



# **Universidad Nacional Mayor de San Marcos**

**Universidad del Perú. Decana de América**

**Facultad de Ciencias Matemáticas**

**Escuela Académico Profesional de Estadística**

## **Identificación de preferencias académicas universitarias en alumnos de los últimos años de educación secundaria en el colegio particular “Bella Unión” mediante el uso del Análisis Conjunto -Perfil Completo- con el aplicativo estadístico R**

### **TESIS**

**Para optar el Título Profesional de Licenciado en Estadística**

### **AUTOR**

**Jhon Frank DE LA CERNA VILLAVICENCIO**

### **ASESOR**

**Rosa Fátima MEDINA MERINO**

**Lima, Perú**

**2016**



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

## Referencia bibliográfica

---

De la Cerna, J. (2016). *Identificación de preferencias académicas universitarias en alumnos de los últimos años de educación secundaria en el colegio particular “Bella Unión” mediante el uso del Análisis Conjunto -Perfil Completo- con el aplicativo estadístico R*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Escuela Académico Profesional de Estadística]. Repositorio institucional Cybertesis UNMSM.

---



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS  
(Universidad del Perú, DECANA DE AMÉRICA)  
FACULTAD DE CIENCIAS MATEMÁTICAS



ESCUELA ACADÉMICO PROFESIONAL DE ESTADÍSTICA

ACTA DE SUSTENTACIÓN DE TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE  
LICENCIADO EN ESTADÍSTICA

En la Ciudad Universitaria, Facultad de Ciencias Matemáticas, siendo las 17:00 horas del día 17 de MAYO del año 2016 se reunieron los docentes designados como miembros del Jurado:

Mg. Olga Lidia Solano Dávila	(Presidente)
Mg. Rosa Ysabel Adriazola Cruz	(Miembro)
Mg. Rosa Fátima Medina Merino	(Miembro Asesor)

Para la sustentación de la Tesis intitulada "IDENTIFICACIÓN DE PREFERENCIAS ACADÉMICAS UNIVERSITARIAS EN ALUMNOS DE LOS ÚLTIMOS AÑOS DE EDUCACIÓN SECUNDARIA EN EL COLEGIO PARTICULAR "BELLA UNIÓN" MEDIANTE EL USO DEL ANÁLISIS CONJUNTO - PERFIL COMPLETO - CON EL APLICATIVO ESTADÍSTICO R", presentada por el Bachiller Jhon Frank De La Cerna Villavicencio, para obtener el Título Profesional de Licenciado en Estadística.

Luego de la exposición de la Tesis, la Presidente invitó al expositor a dar respuesta a las preguntas formuladas.

Realizada la evaluación correspondiente por los miembros del jurado, el expositor mereció la aprobación..... Sobresaliente....., con un calificativo promedio de ..... Diecisiete (17).....(letras y números).

A continuación los miembros del jurado, dan manifiesto que el participante Bachiller Jhon Frank De La Cerna Villavicencio, en virtud de haber aprobado la sustentación de su tesis, será propuesto para que se le otorgue el Título Profesional de Licenciado en Estadística.

Siendo las ..... horas, se levantó la Sesión, firmando para constancia la presente Acta en tres (3) copias originales.

Ysabel Adriazola Cruz  
Mg. Rosa Ysabel Adriazola Cruz  
Miembro

Rosa Fátima Medina Merino  
Mg. Rosa Fátima Medina Merino  
Miembro Asesor

Olga Lidia Solano Dávila  
Mg. Olga Lidia Solano Dávila  
Presidente

**IDENTIFICACIÓN DE PREFERENCIAS ACADÉMICAS UNIVERSITARIAS**  
**EN ALUMNOS DE LOS ÚLTIMOS AÑOS DE EDUCACIÓN SECUNDARIA EN**  
**EL COLEGIO PARTICULAR “BELLA UNIÓN” MEDIANTE EL USO DEL**  
**ANÁLISIS CONJUNTO -PERFIL COMPLETO- CON EL APLICATIVO**  
**ESTADÍSTICO R**

**Jhon Frank de la Cerna Villavicencio**

Tesis presentada a consideración del Cuerpo Docente de la Facultad de Ciencias Matemáticas, de la Universidad Nacional Mayor de San Marcos, como parte de los requisitos para obtener el Título Profesional de Licenciado en Estadística.

Aprobada por:

.....  
Mg. Olga Lidia Solano Dávila  
JURADO PRESIDENTE

.....  
Mg. Rosa Ysabel Adriazola Cruz  
JURADO

.....  
Mg. Rosa Fátima Medina Merino  
JURADO ASESOR

**LIMA - PERÚ**  
**Mayo - 2016**

## **FICHA CATALOGRÁFICA**

**DE LA CERNA VILLAVICENCIO, JHON FRANK**

**Identificación de Preferencias Académicas universitarias en alumnos de los últimos años de educación secundaria en el colegio particular “Bella Unión” mediante el uso del Análisis Conjunto -Perfil Completo- con el aplicativo estadístico R, (Lima) 2016.**

xi, 119p., 29,7cm, (UNMSM, Licenciado, Estadística, 2016)

Tesis, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, 1. Estadística.

I. UNMSM/FdeCM II. Identificación de Preferencias Académicas Universitarias en alumnos de los últimos años de educación secundaria en el colegio particular “Bella Unión” mediante el uso del Análisis Conjunto -Perfil Completo- con el aplicativo estadístico R

## **DEDICATORIA**

**A DIOS, POR SU ETERNA BENDICIÓN. A MI SEÑORA MADRE JULIA VILLAVICENCIO RAMOS, POR SU AMOR Y APOYO INCONDICIONAL, Y A MI ASESORA DE TESIS MG. ROSA FÁTIMA MEDINA MERINO, POR SU ORIENTACIÓN EN LA REALIZACIÓN DE ESTA TESIS.**

## RESUMEN

### **IDENTIFICACIÓN DE PREFERENCIAS ACADÉMICAS UNIVERSITARIAS EN ALUMNOS DE LOS ÚLTIMOS AÑOS DE EDUCACIÓN SECUNDARIA EN EL COLEGIO PARTICULAR “BELLA UNIÓN” MEDIANTE EL USO DEL ANÁLISIS CONJUNTO -PERFIL COMPLETO- CON EL APLICATIVO ESTADÍSTICO R**

**Jhon Frank de la Cerna Villavicencio**

**MAYO - 2016**

Orientador: Mg. Rosa Fátima Medina Merino  
Título obtenido: Licenciado en Estadística

---

El Análisis Conjunto permite conocer y entender las preferencias de las personas acerca de productos y servicios. En este trabajo se investigó sobre esta técnica estadística en la metodología tradicional de perfiles completos, aplicándolo a la identificación de preferencias hacia un “producto”: “Identificación de preferencias universitarias en alumnos de los últimos años de educación secundaria”. Las preferencias universitarias no estaban enfocadas hacia la profesión o carrera universitaria a seguir; sino, a preferencias sobre una determinada institución universitaria considerando aspectos como la calidad de la enseñanza, el tipo de universidad (nacional o particular), la infraestructura académica, etc. El diseño del estudio fue experimental, transversal y explicativo. La selección de los estudiantes se dio a través de un muestreo por selección intencionada. Así mismo, el análisis estadístico se realizó a través del programa estadístico R haciendo uso del package Conjoint.

**PALABRAS CLAVE:**

ANÁLISIS CONJUNTO  
PERFILES COMPLETOS  
ANÁLISIS MULTIVARIADO  
TÉCNICA MULTIVARIANTE  
PACKAGE CONJOINT  
PREFERENCIAS UNIVERSITARIAS



**ABSTRACT****IDENTIFICATION OF UNIVERSITY PREFERENCES IN STUDENTS OF THE  
LAST YEARS OF SECONDARY EDUCATION IN PARTICULAR "BELLA  
UNION" SCHOOL USING TRADITIONAL CONJOINT ANALYSIS OF FULL  
PROFILES WITH R STATISTICAL SOFTWARE****Jhon Frank de la Cerna Villavicencio****MAY - 2016**

Thesis Advisor: Mg. Rosa Fátima Medina Merino  
Academic Degree: Degree in Statistics

---

Conjoint analysis allows to know and understand the preferences of people about products and services. In this work investigated this statistical technique in traditional methodology of full profiles, applying to identify preferences for a “product”: “Identification of university preferences in students of the last years of secondary education”. University preferences were not focused towards the profession or college career to follow; but, preferences on a particular university institution taking into account aspects such as the quality of the education, the type of school (public or private), the academic infrastructure, etc. The design of the study was experimental, cross-sectional, and explanatory. The selection of the students was given through a sampling intentional selection. Likewise, statistical analysis was performed through the statistical software R using Conjoint R package.

**KEYWORDS:**

CONJOINT ANALYSIS  
FULL PROFILES  
MULTIVARIATE ANALYSIS  
MULTIVARIATE TECHNIQUES  
CONJOINT R PACKAGE  
UNIVERSITY PREFERENCES

## TABLA DE CONTENIDOS

INTRODUCCIÓN .....	1
<b>CAPÍTULO I-PLANTEAMIENTO DEL PROBLEMA .....</b>	<b>4</b>
1.1 PROBLEMA DE INVESTIGACIÓN.....	4
1.2 PREGUNTAS DE INVESTIGACIÓN .....	6
1.2.1 PREGUNTA GENERAL .....	6
1.2.2 PREGUNTAS ESPECÍFICAS.....	6
1.3 OBJETIVOS .....	6
1.3.1 OBJETIVO GENERAL .....	6
1.3.2 OBJETIVOS ESPECÍFICOS .....	6
1.4 ANTECEDENTES.....	7
1.4.1 ANTECEDENTES NACIONALES .....	7
1.4.2 ANTECEDENTES INTERNACIONALES .....	10
1.5 JUSTIFICACIÓN E IMPORTANCIA .....	12
1.6 METODOLOGÍA DE TRABAJO.....	13
1.6.1 TIPO DE INVESTIGACIÓN.....	13
1.6.2 DISEÑO DE INVESTIGACIÓN .....	13
1.6.3 POBLACIÓN Y MUESTRA .....	13
1.6.3.1 POBLACIÓN OBJETIVO .....	13
1.6.3.2 UNIDAD DE ANÁLISIS .....	14
1.6.3.3 MUESTRA .....	14
1.6.4 INSTRUMENTOS DE RECOLECCIÓN DE DATOS .....	14
1.6.5 PROCEDIMIENTOS DE RECOLECCIÓN .....	15
1.6.6 IDENTIFICACIÓN Y CLASIFICACIÓN DE LAS VARIABLES.....	16
1.6.6.1 IDENTIFICACIÓN DE LAS VARIABLES .....	16
1.6.6.2 CLASIFICACIÓN DE LAS VARIABLES .....	16
1.6.7 MÉTODOS Y MODELOS DE ANÁLISIS DE DATOS SEGÚN TIPO DE VARIABLES .....	17
1.6.8 TÉCNICAS DE PROCESAMIENTO Y ANÁLISIS DE DATOS .....	17
1.6.9 ASPECTOS ÉTICOS.....	17
<b>CAPÍTULO II-ANÁLISIS CONJUNTO.....</b>	<b>18</b>
2.1 DEFINICIÓN Y FINALIDAD .....	18
2.2 HISTORIA DEL ANÁLISIS CONJUNTO .....	20
2.3 APLICACIÓN DEL ANÁLISIS CONJUNTO.....	22
2.4 ETAPAS DE LA METODOLOGÍA ANÁLISIS CONJUNTO .....	23
2.5 VARIANTES DEL ANÁLISIS CONJUNTO .....	24
2.5.1 ANÁLISIS CONJUNTO DE PERFIL COMPLETO (FULL PROFILE) .....	26
2.5.2 ANÁLISIS CONJUNTO ADAPTATIVO (ACA) .....	27
2.5.3 ANÁLISIS CONJUNTO BASADO EN LA ELECCIÓN .....	29
2.5.4 ANÁLISIS CONJUNTO BASADO EN LA ELECCIÓN DE PERFILES PARCIALES O PARTIAL-PROFILE CBC .....	30
2.5.5 ANÁLISIS CONJUNTO BASADO EN LA ELECCIÓN ADAPTATIVO O ADAPTATIVE CHOICE-BASED CONJOINT (ACBC).....	31
2.6 LA METODOLOGIA CONJOINT .....	32
2.6.1 PLANTEAMIENTO DEL PROBLEMA .....	33

2.6.2 SELECCIÓN DE ATRIBUTOS Y NIVELES .....	34
2.6.3 ESCOGIENDO LA METODOLOGÍA CONJUNTA .....	38
2.6.4 FORMA DE PRESENTACIÓN .....	41
2.6.5 CREACIÓN DE ESTÍMULOS .....	45
2.6.6 TRABAJO DE CAMPO Y TRATAMIENTO DE DATOS .....	50
2.6.7 ESTIMACIÓN DE LAS UTILIDADES .....	51
2.6.8 EVALUACIÓN DEL AJUSTE .....	58
<b>CAPÍTULO III-APLICACIÓN Y RESULTADOS.....</b>	<b>61</b>
3.1 INTRODUCCIÓN .....	61
3.2 PLANTEAMIENTO DEL PROBLEMA .....	62
3.3 SELECCIÓN DE ATRIBUTOS Y NIVELES.....	63
3.4 ESCOGIENDO LA METODOLOGÍA CONJUNTA .....	65
3.5 FORMAS DE PRESENTACIÓN .....	66
3.6 CREACIÓN DE ESTÍMULOS.....	66
3.7 TRABAJO DE CAMPO Y TRATAMIENTO DE DATOS .....	70
3.8 ESTIMACIÓN DE LAS UTILIDADES .....	77
3.9 EVALUACIÓN DEL AJUSTE .....	87
3.10 INTERPRETACIÓN DE LOS RESULTADOS.....	93
3.11 APLICACIONES MÁS COMUNES DE LA REPRESENTACIÓN DE LA ESTRUCTURA DE PREFERENCIAS .....	95
3.11.1 SEGMENTACIÓN .....	95
3.11.2 SIMULACIÓN .....	100
<b>CONCLUSIONES .....</b>	<b>103</b>
<b>LISTA DE REFERENCIAS .....</b>	<b>105</b>
<b>ANEXOS.....</b>	<b>107</b>
ANEXO 1–INSTRUMENTO DE RECOLECCIÓN DE PREFERENCIAS.....	108
ANEXO 2–MATRIZ DE PREFERENCIAS UNIVERSITARIAS .....	109
ANEXO 3–MATRIZ DE PERFILES DE RESERVA (HOLDOUT).....	109
ANEXO 4–MATRIZ DE PERFILES PARA EVALUAR LA FIABILIDAD TEMPORAL....	109
ANEXO 5–MATRIZ DE PREFERENCIAS. PRIMER CONGLOMERADO .....	110
ANEXO 6–MATRIZ DE PREFERENCIAS. SEGUNDO CONGLOMERADO .....	110
ANEXO 7–GLOSARIO DE TÉRMINOS.....	111
ANEXO 8–DESARROLLO MATRICIAL DE LA ESTIMACIÓN OLS CON VARIABLES DUMMY .....	112

## LISTA DE TABLAS

### Capítulo I-PLANTEAMIENTO DEL PROBLEMA

Tabla I.1-Variables de estudio .....	16
--------------------------------------	----

### Capítulo II-ANÁLISIS CONJUNTO

Tabla II.1-Lista de áreas de aplicación del Análisis Conjunto.....	22
Tabla II.2-Etapas del Análisis Conjunto.....	23
Tabla II.3-Principales aspectos distintivos entre el A.C. tradicional y el A.C. basado en la elección .....	25
Tabla II.4-Comparación de metodologías conjuntas alternativas .....	39

### Capítulo III-APLICACIÓN Y RESULTADOS

Tabla III.1-Atributos o factores evaluados inicialmente.....	64
Tabla III.2-Atributos o factores seleccionados .....	64
Tabla III.3-Atributos y niveles seleccionados.....	65
Tabla III.4-Funciones del Conjoint R Package .....	77
Tabla III.5-Estimación utilidades totales del modelo AC .....	91
Tabla III.6-Combinación ideal.....	93
Tabla III.7-Utilidades totales de las tarjetas evaluadas .....	94
Tabla III.8-Comparativo de perfiles ideales.....	100

## LISTA DE FIGURAS

### Capítulo I-PLANTEAMIENTO DEL PROBLEMA

Figura I.1-Ejemplo de un perfil completo para una tarea de cuantificación .....	15
---	----

### Capítulo II-ANÁLISIS CONJUNTO

Figura II.1-Evolución del Análisis Conjunto .....	26
Figura II.2-Atributos de un producto .....	35
Figura II.3-Selección de atributos y niveles: ejemplo viaje de promoción .....	37
Figura II.4-Modelos de preferencia .....	40
Figura II.5-Matriz de comparaciones.....	42
Figura II.6-Ordenación de perfiles completos .....	43
Figura II.7-Cuantificación de perfiles completos.....	43
Figura II.8-Presentación de estímulos en ACA.....	44
Figura II.9-Presentación de estímulos en un Análisis Conjunto basado en la elección .....	45
Figura II.10-Ejemplo de ordenación de perfiles completos .....	51
Figura II.11-Los tres tipos de relaciones básicas entre los niveles de los factores en el AC .....	56

### Capítulo III-APLICACIÓN Y RESULTADOS

Figura III.1-Diseño factorial completo .....	68
Figura III.2-Diseño factorial fraccionado ortogonal .....	69
Figura III.3-Correlaciones entre los atributos .....	70
Figura III.4-Tarjeta N° 1 para el recojo de preferencias .....	70
Figura III.5-Tarjeta N° 2 para el recojo de preferencias .....	71
Figura III.6-Tarjeta N° 3 para el recojo de preferencias .....	71
Figura III.7-Tarjeta N° 4 para el recojo de preferencias .....	71
Figura III.8-Tarjeta N° 5 para el recojo de preferencias .....	72
Figura III.9-Tarjeta N° 6 para el recojo de preferencias .....	72
Figura III.10-Tarjeta N° 7 para el recojo de preferencias .....	72
Figura III.11-Tarjeta N° 8 para el recojo de preferencias .....	73
Figura III.12-Vector con los nombres de los niveles (“level”) .....	74
Figura III.13-Vector de preferencias (“pfe”) .....	75
Figura III.14-Matriz de perfiles codificados (“code”) .....	76
Figura III.15-Matriz de utilidades de los niveles .....	79
Figura III.16-Importancia relativa de los atributos .....	80
Figura III.17-Importancia relativa de los atributos .....	80
Figura III.18-Estimación de las utilidades parciales .....	81
Figura III.19-Tipo de universidad: preferencia de los niveles .....	82
Figura III.20-Reputación académica: preferencia de los niveles .....	82
Figura III.21-Calidad de la enseñanza: preferencia de los niveles .....	83
Figura III.22-Infraestructura académica: preferencia de los niveles .....	83
Figura III.23-Infraestructura deportiva: preferencia de los niveles.....	84
Figura III.24-Perspectiva de empleo: preferencia de los niveles .....	84
Figura III.25-Estimación general .....	85
Figura III.26-Utilidades normalizadas de la evaluación principal .....	87
Figura III.27-Utilidades normalizadas de la evaluación complementaria.....	88
Figura III.28-Estimación individual – encuestado n° 2 .....	90
Figura III.29-Estimación utilidades totales de los perfiles de reserva .....	92
Figura III.30-Estímulo con mayor preferencia.....	94
Figura III.31-Segmentación usando el método k-means.....	95
Figura III.32-Análisis del primer conglomerado .....	97
Figura III.33-Importancia relativa de los atributos en el primer conglomerado .....	97
Figura III.34-Análisis del segundo conglomerado .....	99
Figura III.35-Importancia relativa de los atributos en el segundo conglomerado.....	99
Figura III.36- Matriz de perfiles a simular .....	101
Figura III.37- Simulación de preferencias-porcentajes de participación.....	102

## LISTA DE FÓRMULAS

### Capítulo II-ANÁLISIS CONJUNTO

Fórmula II.1-Número de matrices de comparaciones .....	45
Fórmula II.2-Número de combinaciones en el método de perfiles completos.....	46
Fórmula II.3-Número mínimo de estímulos .....	49
Fórmula II.4-Modelo aditivo lineal.....	53
Fórmula II.5-Modelo básico de regresión.....	54
Fórmula II.6-Modelo de preferencias globales .....	54
Fórmula II.7-Modelo de preferencias en su forma matricial .....	57
Fórmula II.8-Importancia relativa de un atributo.....	58

## INTRODUCCIÓN

En diferentes investigaciones surge la necesidad de medir las preferencias que tienen los individuos hacia determinados productos y servicios. Estos análisis son requeridos en diferentes áreas y principalmente a nivel del Marketing. Las técnicas estadísticas utilizadas para este fin van desde los métodos descriptivos hasta los multivariados.

El presente estudio abarca una de las técnicas estadísticas que permiten evaluar preferencias: el Análisis Conjunto. El Análisis Conjunto es una técnica estadística multivariante que permite conocer y entender las preferencias de las personas hacia los productos y servicios. También, se la podría definir, de una forma resumida y clara, como lo señaló el Mg. Jesús Salinas Flores en el XVI CONEEST<sup>1</sup> desarrollado en la Universidad Nacional Mayor de San Marcos, “el Análisis Conjunto es una herramienta estadística que permite medir cuál es la opción que más le agrada al cliente, teniendo en cuenta todos los atributos y niveles que puede tener el producto o servicio ofertado” (2012, p. 2)

Dentro de los tipos de metodologías que actualmente presenta el Análisis Conjunto podemos mencionar: “el Análisis Conjunto tradicional y el Análisis de Elección Experimental, conocido también como Análisis Conjunto Basado en Elección” (Picón, Varela, & Braña, 2006, p. 17). La presente investigación se enfoca en el estudio del Análisis Conjunto tradicional o clásico de perfiles completos.

Además de desarrollar la teoría de esta técnica, este trabajo desarrolla una aplicación del Análisis Conjunto tradicional de perfiles completos dirigido a la identificación de

---

<sup>1</sup> Congreso Nacional de Estudiantes de Estadística

preferencias hacia un “producto”: la identificación de preferencias universitarias. El cual constituye su objetivo principal.

Estas preferencias universitarias no buscan conocer la profesión o carrera universitaria a seguir; sino, se fija en otras características, de igual o mayor importancia, que determinan la elección hacia una determinada universidad. Por ejemplo: el tipo de universidad (nacional o particular), el prestigio o reputación académica, la infraestructura académica, la perspectiva de empleo, etc.

Otro propósito de la investigación es el desarrollo del análisis estadístico utilizando el programa estadístico R; a través, de las funciones del package Conjoint cuyos algoritmos están implementados bajo el Análisis Conjunto tradicional de perfiles completos y el modelo de regresión lineal múltiple con variables ficticias (Bak & Bartlomowicz, 2009).

Así mismo, otra finalidad del presente estudio, es detallar las aplicaciones de la representación de la estructura de preferencias. Las aplicaciones más comunes incluyen la segmentación, el análisis de renta y simuladores conjuntos (Hair, Anderson, Tatham, & Black, 1999).

El diseño de la presente investigación es experimental, transversal y explicativo. La población de estudio está comprendido por estudiantes de los dos últimos años de educación secundaria (4to. y 5to. secundaria), seleccionados a conveniencia, de un colegio particular ubicado en el distrito Callao de la Provincia Constitucional del mismo nombre.



La composición del documento se divide en tres capítulos. En el Capítulo I-Planteamiento del problema, se muestra las secciones referidas al problema de investigación, las preguntas de investigación, los objetivos de estudio, los antecedentes, la justificación e importancia del estudio y la metodología de trabajo.

En el Capítulo II-Análisis Conjunto, se presenta la teoría concerniente a la técnica Análisis Conjunto. Si bien se describe todas las metodologías del Análisis Conjunto, se prioriza los conceptos hacia el método tradicional o clásico.

Por último, en el Capítulo III-Aplicación y Resultados, se desarrolla la aplicación de la técnica Análisis Conjunto tradicional en la identificación de preferencias universitarias. Bajo el criterio de perfiles completos. Detallando los pasos realizados a través de la metodología conjunta y el uso de las funciones del paquete Conjoint en el programa estadístico R.

# **CAPÍTULO I**

## **PLANTEAMIENTO DEL PROBLEMA**

### **1.1 PROBLEMA DE INVESTIGACIÓN**

Una primera decisión al realizar estudios sobre las preferencias de las personas es evaluar la metodología a considerar para el análisis de éstas.

Según Varela, Rial y Picón (2014), tenemos, principalmente, 3 tipos de métodos: holísticos, composicionales y descomposicionales (multiatributo). En el análisis holístico, se desconoce la estructura de las preferencias; mientras, en el análisis composicional el estudio de cada atributo (cualidad o característica del producto o servicio) se da por separado. Por último, el análisis multiatributo, se distingue de las demás por ser la más realista, ya que el individuo tiene que valorar todas las características a la vez antes de elegir un producto o servicio.

El Análisis Conjunto (AC) es una técnica estadística multivariante que “se define como descomposicional y multiatributiva” (Varela et al., 2014, p. 9). Debido a que “estima la estructura de las preferencias de un consumidor (por ejemplo: mediante utilidades

parciales, ponderaciones de importancia, puntos ideales, etc.), una vez obtenidas sus evaluaciones globales de un conjunto de alternativas preespecificadas en términos de niveles de diferentes atributos” (Green & Srinivasan, 1978, p. 104).

Esta técnica se condice con la manera en que normalmente las personas eligen los productos y servicios. Por ejemplo, si se desea adquirir un nuevo equipo celular, es común evaluar la marca (Motorola, Samsung, Apple, etc.), el tamaño (pequeño, mediano, grande), el color (blanco, rojo, etc.), el plan tarifario (prepago, pospago), etc. Es decir, se valoran diferentes atributos con sus respectivos niveles a la vez antes de tomar una decisión.

Actualmente dentro del Análisis Conjunto coexisten dos metodologías conjuntas: el tradicional o clásico y el basado en la elección. Estas a su vez han evolucionado en nuevas metodologías como el Análisis Conjunto Adaptativo (ACA), El Análisis Conjunto Adaptativo basado en la Elección, etc.

La presente investigación utiliza y presenta al Análisis Conjunto tradicional de perfiles completos como una de las mejores opciones al momento de identificar preferencias de los consumidores hacia un producto o servicio. Desarrollando, con fines de un mayor entendimiento de la técnica, una aplicación orientada a la identificación de preferencias hacia un “producto”: la identificación de las preferencias al escoger una determinada universidad por parte de estudiantes de los últimos dos años en etapa escolar. También propone, el uso del programa estadístico R para el análisis del AC. Ya que no es muy usado en el análisis de esta técnica. R es un programa libre y de fácil acceso compatible con diferentes sistemas operativos.

## **1.2 PREGUNTAS DE INVESTIGACIÓN**

### **1.2.1 PREGUNTA GENERAL**

¿Cómo se desarrolla la técnica Análisis Conjunto tradicional de perfiles completos empleando el caso práctico “identificación preferencias universitarias en alumnos de los últimos años de educación secundaria”?

### **1.2.2 PREGUNTAS ESPECÍFICAS**

- ¿Cómo se elabora el análisis estadístico, utilizando el programa R, de la técnica Análisis Conjunto tradicional de perfiles completos en el caso práctico “identificación de preferencias universitarias en alumnos de los últimos años de educación secundaria”?
- ¿Cuáles son las aplicaciones más comunes que se dan a la representación de la estructura de preferencias del Análisis Conjunto tradicional de perfiles completos obtenidos en la identificación de preferencias universitarias en alumnos de los últimos años de educación secundaria?

## **1.3 OBJETIVOS**

### **1.3.1 OBJETIVO GENERAL**

Explicar la técnica Análisis Conjunto tradicional de perfiles completos, empleando para este fin, un caso práctico, la “identificación de preferencias universitarias en alumnos de los últimos años de educación secundaria”.

### **1.3.2 OBJETIVOS ESPECÍFICOS**

- Elaborar el análisis estadístico, utilizando el programa estadístico R, en el desarrollo de la técnica Análisis Conjunto tradicional de perfiles completos para la

“identificación de preferencias universitarias en alumnos de los últimos años de educación secundaria”.

- Desarrollar las aplicaciones más comunes de la representación de la estructura de preferencias del Análisis Conjunto tradicional de perfiles completos obtenidos en la “identificación de preferencias universitarias en alumnos de los últimos años de educación secundaria”.

## 1.4 ANTECEDENTES

En este apartado procederemos a mencionar las investigaciones, tanto a nivel nacional e internacional, realizadas aplicando el Análisis Conjunto tradicional de perfiles completos. Y mucho mejor si éstas han sido realizadas enfocados en el tema de estudio: la identificación de preferencias universitarias.

### 1.4.1 ANTECEDENTES NACIONALES

A nivel nacional, son escasas las investigaciones realizadas aplicando el AC tradicional y aún más, si se trata de identificar preferencias universitarias. No obstante, se pudo identificar algunos trabajos que se mencionan a continuación:

La tesis realizada por Edgar Cárdenas (2006) con el tema **“EN BUSCA DEL PERFIL DEL CANDIDATO POLÍTICO GANADOR MEDIANTE LA APLICACIÓN DEL ANÁLISIS CONJUNTO”**, investigación que tiene como objetivo principal desarrollar un modelo de la estructura de decisión del elector al momento de elegir al candidato de su preferencia. Esto es, asumiendo que la estructura del proceso de decisión de voto del elector en un proceso electoral, presenta las mismas componentes

estructurales que el proceso de toma de decisiones del consumidor al momento de elegir un producto o servicio. Esto es, considerando al candidato político como un producto ‘vendible’ en el mercado, por lo tanto presentará atributos y niveles. Para este fin utiliza la técnica **Análisis Conjunto tradicional de perfiles completos**. Luego, el análisis estadístico lo desarrolla utilizando el programa SPSS con sus módulos ORTHOPLAN para generar diseños ortogonales, PLANCARDS para producir las tarjetas con los perfiles completos y el CONJOINT, el cual permite proporcionar información de cómo los encuestados piensan y cómo valoran los distintos atributos.

El trabajo desarrollado por Matthew Bird, Liuba Kogan, Luz Sánchez y Julio Villa (2014) “**HACIA UN NUEVO MODELO DE LAS DECISIONES DE COMPRA: LÓGICAS DE CONSUMO EN JÓVENES UNIVERSITARIOS LIMEÑOS**”. Los autores señalan que: el estudio toma el caso peruano para elaborar un modelo novedoso para las decisiones de compra y para que en el futuro contribuya a comprender y analizar a los consumidores globales salientes. Por lo tanto, plantean como objetivo principal proponer un novedoso modelo que permita comprender las lógicas de consumo de los jóvenes universitarios para poder predecir sus preferencias. Además, señalan que este modelo de decisión de consumo se elaboró en tres fases: primero, identificación de la lógica del consumo. Segundo validar las lógicas y desarrollar los perfiles de los universitarios, y por último, en base a esos perfiles, modelar su toma de decisión. La metodología conjunta utilizada fue el **Análisis Conjunto Basado en la Elección**. El análisis se realizó utilizando el programa PQMethod.

El siguiente estudio, a nivel nacional, es el desarrollado por Percy Marquina y Arturo Vásquez (2013), “**CONSUMER SOCIAL RESPONSES TO CSR INITIATIVES**

**VERSUS CORPORATE ABILITIES**". Los autores indican que: "las respuestas del consumidor a las iniciativas de Responsabilidad Social Empresarial (RSE) en comparación con las respuestas del consumidor a Habilidades Corporativas (HC) han sido esquivos. Las investigaciones empíricas pertinentes sobre el tema muestran resultados poco claros". Por lo tanto, proponen como objetivo principal de esta investigación examinar los antecedentes claves a las respuestas sociales del consumidor, en forma particular, los efectos que se desprenden de comparar las iniciativas del RSE y HC en el comportamiento de compra de los consumidores. Para este fin se aplicó el **Análisis Conjunto Basado en la Elección** a muestras por cuotas del consumo de dos países diferentes (Estados Unidos y Perú) en la industria del calzado.

Por último, la investigación "**CARACTERÍSTICAS ESTEREOTÍPICAS QUE INTERVIENEN EN EL PROCESO DE EVALUACIÓN DE CANDIDATOS ELECTORALES**" realizada por Korina Eiko Papuico Navarro (2015). El estudio tiene por objetivo principal identificar y analizar la importancia relativa que tienen las características estereotípicas, esto es sexo, raza y edad, en las preferencias hacia un candidato presidencial en los universitarios de Lima Metropolitana. Además, se trabaja con la **metodología Análisis Conjunto basado en la Elección**, utilizando el paquete estadístico CBC ® (Choice-Based Conjoint) 4.0 de Sawtooth Software para el análisis estadístico.

Como se observa, la mayoría de estudios mencionados a nivel nacional utilizan el Análisis Conjunto Basado en la Elección. Y, en ninguno de los casos se utiliza el programa estadístico R.

#### 1.4.2 ANTECEDENTES INTERNACIONALES

A continuación se presentan los antecedentes del uso de la técnica Análisis Conjunto a nivel internacional.

El estudio realizado por Geoffrey Soutar y Julia Turner (2002), **“STUDENTS' PREFERENCES FOR UNIVERSITY: A CONJOINT ANALYSIS”**. Donde se examinan las preferencias universitarias de egresados de la escuela secundaria en Australia. Investigando la importancia de una serie de atributos considerados (tipo de curso, reputación académica, “atmósfera” del campus universitario, calidad de la enseñanza, tipo de universidad, distancia desde la casa, etc.). Para este fin se trabaja con una variante del Análisis Conjunto “clásico”, la metodología **Análisis Conjunto Adaptativo (ACA)**. Se utilizó el programa SPSS para el análisis estadístico y dentro de los resultados obtenidos más relevantes podemos mencionar los atributos que presentaron un mayor nivel de importancia relativa: la idoneidad del curso (15%), la reputación académica (12%) y las perspectivas de empleo (12%).

La investigación efectuada por Caroline Walsh, Cornelius Barry, Andrew Dunnett y Jan Moorhouse (2013), **“CHOOSING A UNIVERSITY: THE RESULTS OF A LONGITUDINAL STUDY USING CONJOINT ANALYSIS”**. Donde se toma como antecedente y referencia el estudio realizado por Dunnett et al. (2012), donde se observó que los cursos y la reputación de la universidad fueron los atributos más importantes que influyen en la elección sobre una determinada universidad. Por lo tanto, se llevó a cabo un estudio longitudinal con la misma muestra de estudiantes. Concentrándose en los encuestados que habían hecho una opción real, la intención era abordar las deficiencias del anterior estudio y evaluar el análisis conjunto como una herramienta



con poder predictivo. Por lo tanto, los objetivos del estudio se centraron: en primer lugar, evaluar la importancia de los factores que afectan a los estudiantes al escoger una universidad, y el impacto en diversos segmentos del mercado estudiantil. En segundo lugar, evaluar si el método Conjoint puede ser visto como un predictor fiable de la importancia de los atributos que afectan las preferencias. Los atributos considerados fueron: reputación de la universidad, reputación de los cursos, requisitos de ingreso, las cuotas anuales, orientación de la universidad y distancia desde la casa. No se especifica la metodología del AC aplicado pero se deduce de la revisión que es el **Análisis Conjunto tradicional**. Los resultados dan cuenta de los atributos que obtuvieron un mayor nivel de importancia relativa: la reputación de los cursos (31,2%) y la reputación de la universidad (26,2%).

La investigación realizada por Jan Moorhouse (2012), “WHAT PRICE UNIVERSITY EDUCATION?: EVIDENCE FROM A CONJOINT ANALYSIS”. Donde el estudio parte de la problemática que representaría el incremento de las tarifas en los cursos que brindan las universidades de Inglaterra y la necesidad de éstas en realizar investigaciones del impacto de las tasas mayores en la decisión de los estudiantes. A su vez, el autor señala que su estudio utiliza al Análisis Conjunto para simular opciones de preferencia de los estudiantes. Para este fin, se utiliza una serie de criterios que afectan la elección del curso en la Universidad, basado en la literatura existente. Los resultados indican que la Reputación del Curso y la Reputación de la Universidad son mucho más importantes que los Honorarios. En este trabajo se aplica el **Análisis Conjunto tradicional**.

En este casos, si se pudo detectar aplicaciones del Análisis Conjunto tradicional y enfocados a la identificación de preferencias universitarias. No obstante, ninguno de los estudios utilizó el programa estadístico R para llevar a cabo los análisis

## **1.5 JUSTIFICACIÓN E IMPORTANCIA**

Las razones, contribuciones e impacto del desarrollo del presente estudio responden a diferentes aspectos que se mencionan en las siguientes líneas.

Desde el punto de vista teórico, el presente trabajo contribuye en la aplicación del Análisis Conjunto, a nivel de metodología perfil completo. Técnica estadística cuyos orígenes se remontan al trabajo de Luce y Tuckey (1964). Además, se desarrolla un caso práctico, de manera didáctica, sobre identificación de preferencias hacia un “producto”. Con fines de ser replicado en investigaciones de diferentes áreas y no solo a nivel del Marketing. Esto es, en estudios donde de la valoración de los individuos a un conjunto de alternativas, se pretenda deducir la importancia de cada uno de sus atributos y niveles.

Desde el punto de vista metodológico, el presente estudio contribuye al aplicar el Análisis Conjunto como un método de investigación compuesta de varias etapas. Donde algunos autores señalan como una primera etapa a la selección del modelo de preferencia (Green & Srinivasan, 1978), otros autores desde la identificación de atributos y niveles determinantes (Juez & Díez, 1996; Pérez, 2011) o también desde la definición de objetivos (Picón et al., 2006) o planteamiento del problema (Varela et al., 2014)

Desde el punto de vista práctico, el presente estudio contribuye al desarrollar la aplicación del Análisis Conjunto tradicional de perfiles completos haciendo uso del programa estadístico R a través del Conjoint R package (Bak & Bartlomowicz, 2009). El aplicativo R es poco utilizado en el análisis de la técnica AC. Tal como se pudo observar a nivel de los antecedentes nacionales e internacionales. El uso del programa R no demanda costo. Esto debido a que es un aplicativo libre y de fácil descarga<sup>2</sup>. El hecho que sea libre implica la no compra de una licencia para su uso. Otra ventaja de este programa, es el de poder trabajar a nivel de sistemas operativos con diferentes arquitecturas (multiplataforma).

## **1.6 METODOLOGÍA DE TRABAJO**

### **1.6.1 TIPO DE INVESTIGACIÓN**

El tipo de investigación realizado, según los propósitos planteados, es la Investigación Aplicada.

### **1.6.2 DISEÑO DE INVESTIGACIÓN**

El diseño de investigación es experimental, transversal y explicativo.

### **1.6.3 POBLACIÓN Y MUESTRA**

#### **1.6.3.1 POBLACIÓN OBJETIVO**

Estudiantes de los dos últimos años de educación secundaria (4to. y 5to. de secundaria) de los colegios ubicados en la Provincia Constitucional del Callao en el año 2014.

---

<sup>2</sup> <https://cran.r-project.org/>

### **1.6.3.2 UNIDAD DE ANÁLISIS**

Un estudiante de los dos últimos años de educación secundaria (4to. y 5to.) de los colegios ubicados en la Provincia Constitucional del Callao en el año 2014.

### **1.6.3.3 MUESTRA**

La muestra estuvo conformada por 50 estudiantes de los niveles 4to. y 5to. de secundaria de la IEP Bella Unión ubicada en el distrito del Callao, Provincia Constitucional del Callao, en el 2014. El criterio de selección se dio a través de un **Muestreo por Selección Intencionada**. El cual consiste en la elección por métodos no aleatorios de una muestra con características similares a las de la población objetivo.

La obtención de la muestra en función al criterio de selección antes mencionado responde a dos aspectos fundamentales propios de la presente investigación. El primero, el estudio tiene como objetivo principal el explicar la técnica Análisis Conjunto tradicional de perfil completo. Para lograr este propósito y con fines de ser mejor comprendido se desarrolló el caso práctico: “identificación de preferencias universitarias en alumnos de los últimos años de educación secundaria”.

En segundo lugar, el uso del método considerado (Análisis Conjunto tradicional de perfil completo) implicaba un nivel de análisis individual (Hair et al., 1999). Esto es, la identificación de preferencias por cada individuo.

### **1.6.4 INSTRUMENTOS DE RECOLECCIÓN DE DATOS**

Se empleó un instrumento de recolección de datos (ver Anexo 1) compuesto por tarjetas con diversos perfiles estructurados en función de los atributos y niveles considerados

para el estudio. La presentación de las tarjetas fue desarrollada de acuerdo al criterio de Perfiles Completos para el recojo de datos y diseñadas según la forma propuesta por Picón et al. (2006). Un ejemplo de estas tarjetas lo observamos en la figura (I.1).

Figura I.1-Ejemplo de un perfil completo para una tarea de cuantificación

Estudio sobre destinos turísticos											
Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.											
ENTORNO						Naturaleza					
OFERTA GASTRONÓMICA						Normal					
COSTE DEL VIAJE						Medio					
OFERTA LÚDICA						Excelente					
TIPO DE ALOJAMIENTO						Extrahotelero					
OFERTA CULTURAL						Normal					
1	2	3	4	5	6	7	8	9	10		
Muy poco preferida						Extremadamente preferida					

FUENTE: Picón et al. (2006)

### 1.6.5 PROCEDIMIENTOS DE RECOLECCIÓN

A nivel de un estudio Conjunto existen diferentes modos de recoger la información. El cual está estrechamente relacionado con la metodología empleada del AC. Para el presente estudio, donde se aplica el Análisis Conjunto tradicional, se utiliza uno de los criterios, presentando tarjetas donde se les solicita a los encuestados que expresen numéricamente (puntuaciones) el grado de preferencia por cada uno de los perfiles evaluados.

El recojo de los datos se dio a través de una entrevista personal donde se hizo entrega, a cada alumno, de una hoja que contenía las tarjetas estructuradas sobre preferencias universitarias. El plazo de llenado de las cartillas –media hora- se determinó en función de la cantidad de tarjetas a llenar (ocho en total). Así mismo, se verificó el correcto llenado de la hojas al momento del recojo.

## 1.6.6 IDENTIFICACIÓN Y CLASIFICACIÓN DE LAS VARIABLES

### 1.6.6.1 IDENTIFICACIÓN DE LAS VARIABLES

#### ○ Variable dependiente (Y: Preferencia Universitaria)

La Preferencia Universitaria se define como la asignación de una puntuación de preferencia a cada perfil. El intervalo establecido para la aplicación del presente estudio: identificación de preferencias universitarias evaluados por los estudiantes de educación secundaria es del 1 al 10 (mayores valores indican una mayor preferencia).

#### ○ Variables independientes (X<sub>ij</sub>: Niveles del Atributo)

Las variables independientes: Niveles del Atributo, toman valores de presencia (1) o ausencia (0) por cada atributo presentado en los distintos perfiles de identificación de preferencias universitarias.

### 1.6.6.2 CLASIFICACIÓN DE LAS VARIABLES

Las variables de estudio se clasificaron bajo los criterios mostrados en la tabla (I.1)

Tabla I.1-Variables de estudio

VARIABLE	NATURALEZA	TIPO	CRITERIO DE MEDICIÓN	ESCALA
<b>Edad</b>	Cuantitativa	Continua	Edad en años	Razón
<b>Sexo</b>	Cualitativa	Dicotómica	0: Femenino 1: Masculino	Nominal
<b>Nivel de estudios</b>	Cualitativa	Dicotómica	1: 4to. Secundaria 2: 5to. Secundaria	Nominal
<b>Y: Preferencia Universitaria</b>	Cuantitativa	Discreta	Puntuaciones del 1 al 10	Intervalo
<b>X: Niveles del Atributo</b>	Cualitativa	Dicotómica	0: Ausencia 1: Presencia	Nominal

FUENTE: Elaboración Propia

### **1.6.7 MÉTODOS Y MODELOS DE ANÁLISIS DE DATOS SEGÚN TIPO DE VARIABLES.**

Se emplea el Análisis Multivariante haciendo uso de la técnica Análisis Conjunto en su metodología tradicional de perfiles completos.

### **1.6.8 TÉCNICAS DE PROCESAMIENTO Y ANÁLISIS DE DATOS**

Los valores obtenidos en los puntajes de preferencias universitarias, los niveles de los atributos, etc. se tabularon para su posterior análisis a través del programa estadístico R v. 3.1.1.

### **1.6.9 ASPECTOS ÉTICOS**

Para la realización del estudio se contó con el permiso de la dirección de la Institución Educativa Particular representada por la Lic. Hilda Guevara Sarmiento. Además, los alumnos considerados en la investigación colaboraron voluntariamente; previa consulta a los padres de familia quienes mostraron su conformidad y aprobación, en las reuniones convocadas para este fin.

# CAPÍTULO II

## ANÁLISIS CONJUNTO

### 2.1 DEFINICIÓN Y FINALIDAD

El Análisis Conjunto es una técnica aplicable a diferentes áreas del conocimiento: psicología, economía, etc. Sus orígenes se remontan al trabajo de Luce y Tuckey (1964), y a lo largo de la historia ha sabido ganarse un espacio no solo como una técnica; sino incluso, como una metodología de investigación.

De la literatura revisada, mencionaremos algunas de las definiciones que dan respuesta a la pregunta **¿Qué es el Análisis Conjunto?**

Cualquier método descomposicional que estima la estructura de las preferencias de un consumidor (por ejemplo: mediante utilidades parciales, ponderaciones de importancia, puntos ideales, etc.), una vez obtenidas sus evaluaciones globales de un conjunto de alternativas preespecificadas en términos de niveles de diferentes atributos. (Green & Srinivasan, 1978, p. 104)



El Análisis Conjunto es una técnica multivariante que se utiliza específicamente para entender cómo los encuestados desarrollan preferencias acerca de productos y servicios. Se basa en la simple premisa de que los consumidores evalúan el valor de un producto/servicio/idea (real o hipotético) combinando cantidades separadas de valor que proporciona cada atributo. (Hair et al., 1999, p. 408)

El Análisis Conjunto es una técnica estadística de la dependencia utilizada para analizar la relación lineal o no lineal entre una variable dependiente (o endógena) generalmente ordinal (aunque también puede ser métrica) y varias variables independientes (o exógenas) no métricas. (Pérez, 2011, p. 445)

El Análisis Conjunto constituye una técnica multivariante sencilla, potente y flexible para evaluar y analizar las preferencias manifiestas de los consumidores por productos y servicios. La base conceptual para medir las preferencias de los sujetos es la utilidad. (Picón et al., 2006, p. 9)

Luego, con respecto a la finalidad del uso de la técnica Análisis Conjunto, se verifica que existen dos finalidades básicas:

Determinar la contribución de las variables predictoras (atributos) y sus niveles en las preferencias globales del consumidor [...].

Establecer un modelo válido de los juicios de preferencia del consumidor. Los modelos válidos nos permiten predecir la aceptación por parte del consumidor de cualquier combinación de atributos, incluso de aquellos no originariamente evaluados por los consumidores. (Hair et al., 1999, p. 419)

De estas dos finalidades básicas se desprenden otros puntos que se consideran también muy importantes y que justifican el uso del Análisis Conjunto.

Definir el producto o servicio con la mejor combinación posible de niveles de atributos [...], identificar oportunidades de mercado mediante la exploración del potencial de productos o servicios no existentes actualmente [...], detectar el precio óptimo para un producto concreto real o hipotético [...], predecir las cuotas de mercado de diferentes producto con diferentes características [...], detectar grupos de consumidores diferenciados respecto a la estructura de sus preferencias. (Picón et al., 2006, pp. 13-14)

## **2.2 HISTORIA DEL ANÁLISIS CONJUNTO**

En este subcapítulo vamos a realizar una reseña histórica de la técnica Análisis Conjunto desde sus orígenes hasta nuestros días.

Existe consenso en señalar que los orígenes del Análisis Conjunto se remontan al trabajo realizado por los investigadores R. Duncan Luce y John W. Tukey; a través de

su artículo: “Simultaneous conjoint measurement: A new type of fundamental measurement” publicado en 1964. Luce, sicólogo matemático, y Tukey, estadístico, realizan su publicación en el Journal of Mathematical Psychology.

Siguieron a este estudio otras investigaciones y aplicaciones principalmente en el campo de la sicología matemática. Hasta a inicios de los 70, donde los profesores de marketing Paul Green y Vithala Rao adaptaron las Mediciones Conjuntas Simultáneas (Luce & Tukey, 1964) hacia el campo de la investigación comercial. En su artículo “Conjoint Measurement for Quantifying Judgmental Data” señalan que “la medida conjunta es un nuevo desarrollo en Psicología matemática que puede usarse para medir los efectos conjuntos de un conjunto de variables independientes en el orden de una variable dependiente” (1971, p. 1).

El origen del vocablo **Análisis Conjunto** se remonta a 1978, cuando Green y Srinivasan lo utilizan, en el artículo “Conjoint analysis in consumer research: issues and Outlook”, para describir las preferencias del consumidor en el terreno del comportamiento del consumidor (Varela, Rial, & García, 2003).

Desde entonces se vio un interés, años tras año, en la aplicación de la técnica Análisis Conjunto a nivel de los investigadores de mercado. Así lo corroboran diversos artículos donde se presentan resultados de encuestas realizadas a empresas de consultoría especializadas en marketing. Por ejemplo, los realizados por los docentes de marketing Dick R. Wittink y Philippe Cattin (1982, 1989) en Estados Unidos.

Ya en la década del 90, el uso del Análisis Conjunto se incrementó abarcando diferentes áreas de estudio. Y no solo a nivel de los Estados Unidos. Incluso, en forma paralela otras partes del mundo, principalmente en Europa. Así mismo, se elaboraron métodos alternativos de la construcción de las elecciones para los consumidores y de estimación de los modelos conjuntos (Hair et al., 1999).

La implementación de nuevas metodologías del Análisis Conjunto se ha dado, en gran parte, por el desarrollo de los programas informáticos. Los aplicativos permiten una automatización de todo el proceso Conjunto. Un ejemplo, es la metodología que se desprende del Análisis Conjunto tradicional: el Análisis Conjunto Adaptativo (ACA). El cual surge como un programa informático elaborado por Johnson en 1987 a través de su compañía Sawtooth Software.

## 2.3 APLICACIÓN DEL ANÁLISIS CONJUNTO

El Análisis Conjunto, desde sus inicios, se ha sido empleando en proyectos de investigación de marketing aplicadas a diferentes áreas. Una muestra de ello lo observamos en la tabla (II.1).

Tabla II.1-Lista de áreas de aplicación del Análisis Conjunto

Bienes de consumo no duraderos	Productos industriales	Otros productos
1. Jabones de barra	1. Máquinas copadoras	1. Diseño de automóviles
2. Champús para el cabello	2. Equipos de impresión	2. Neumáticos para automóviles y camiones
3. Limpiadores de alfombras	3. Transmisiones de fax	3. Baterías de coches
4. Prendas de fibra sintética	4. Transmisión de datos	4. Medicamentos
5. Precios de Gasolina	5. Terminales informáticos portátiles	5. Tostadora / hornos
6. Pantimedias	6. Diseño computadora personal	6. Cámaras
7. Productos químicos de césped		7. Diseño de apartamento
Servicios financieros	Transporte	Otros servicios
1. Servicios de bancos	1. Aerolíneas nacionales	1. Agencias de alquiler de coches
2. Auto seguros	2. Aerolíneas transcontinentales	2. Servicios telefónicos y fijación de precios
3. Seguros de salud	3. Operaciones del tren de pasajero	3. Agencias de empleo
4. Características de tarjetas de crédito	4. Operaciones del tren de carga	4. Servicios de recuperación de información
5. Tarjetas de descuento de consumo	5. Asociación de transporte aéreo internacional	5. laboratorios
6. Facilidades en venta de autos	6. Diseño de carros electrónicos	6. Diseño de hoteles
7. Servicios de mantenimiento de alta tecnología		

FUENTE: Rao (2014)

Así mismo, "el Análisis Conjunto está siendo utilizado con mayor frecuencia en el ámbito turístico, obteniendo resultados interesantes en lo que respecta a la segmentación de mercados, al conocimiento de las preferencias de los consumidores y a la gestión de estrategias de Marketing-mix" (Lopes, 2011, p. 361). En razón de las diversas variantes o metodologías del Análisis Conjunto.

## 2.4 ETAPAS DE LA METODOLOGÍA ANÁLISIS CONJUNTO

El Análisis Conjunto es concebido no solo como una técnica, sino incluso como una metodología de investigación que incluye una serie de pasos a cumplir.

Al respecto, podemos citar a los investigadores Paul E. Green y V. Srinivasan (1978, 1990) quienes proponen una serie de etapas a realizarse cuando se aplica el Análisis Conjunto, se pueden observar estas etapas en la tabla (II.2).

Tabla II.2-Etapas del Análisis Conjunto

Pasos	Métodos alternativos
1. Selección del modelo de preferencia	Modelo vectorial, Modelo de punto ideal, Modelo de pesos parciales, Modelos mixtos
2. Método de recogida de datos	Dos factores a la vez (matrices trade-off), Perfil completo (evaluación de concepto)
3. Construcción de estímulos para el método de perfiles completos	Diseño Factorial fraccionado, Muestreo aleatorio de distribución multivariante
4. Presentación de los estímulos	Descripción verbal (múltiples señales, tarjeta del estímulo), párrafos descriptivos, representación del modelo pictórico o tridimensional
5. Escala de medición de la variable dependiente	Comparaciones pareadas, orden de rango, escalas de calificación, suma constante de comparaciones pareadas, asignación de la categoría
6. Método de estimación	MONANOVA, PREFMAP, LINMAP, algoritmo de compensación métrico de Johnson, regresión múltiple, LOGIT, PROBIT

FUENTE: Green and Srinivasan (1978)

A todas estas etapas mencionadas, habría que añadirle una primera etapa, la etapa 0 de elección de atributos y niveles determinantes (Juez & Díez, 1996; Pérez, 2011). Si tenemos como objetivo el determinar la contribución de cada atributo y sus niveles en las preferencias globales, resulta de suma importancia esta etapa que consiste en saber escoger los atributos y niveles que formarán parte del estudio.

## **2.5 VARIANTES DEL ANÁLISIS CONJUNTO**

Las metodologías del análisis Conjunto determinan la forma como se mide y analiza las preferencias de los individuos.

Según Picón et al. (2006) en la actualidad coexisten dos grandes paradigmas conjuntos, estrechamente relacionados, aunque manteniendo notables diferencias teóricas y metodológicas, entre ellos: el Análisis Conjunto tradicional y el Análisis de elección experimental o conocido también como Análisis Conjunto Basado en la Elección (Louviere, 1988; Louviere y Woodworth, 1983). (p. 16)

Estas dos variantes se diferencian en el modo del recojo de los datos. Mientras en el Análisis Conjunto tradicional de perfiles completos, los encuestados deben ordenar o asignar una puntuación a cada perfil revisado. En el Análisis Conjunto Basado en la Elección los individuos deben escoger su alternativa preferida de entre un menor número de perfiles. Pudiendo incluso, como sucede en la realidad, el no escoger ninguno (Picón et al., 2006).

En la tabla (II.3) observaremos las principales características que diferencian a cada una de las metodologías del Análisis Conjunto antes mencionadas.

Tabla II.3- Principales aspectos distintivos entre el Análisis Conjunto tradicional y el Análisis Conjunto Basado en la Elección

	<b>Análisis Conjunto tradicional</b> (perfiles completos)	<b>Análisis Conjunto Basado en la Elección</b>
<b>Base teórica comportamental</b>	?	RUT
<b>Método de recogida de datos</b>	Ordenación o puntuación de perfiles	Elección de la alternativa más preferida
<b>Forma del modelo</b>	Aditivo (efectos principales)	Aditivo + interacciones
<b>Método de estimación</b>	Regresión OLS	Logit o Probit Multinomial
<b>datos producidos</b>	A nivel individual	A nivel agregado (Bayesiano a nivel individual)

FUENTE: Picón et al. (2006)

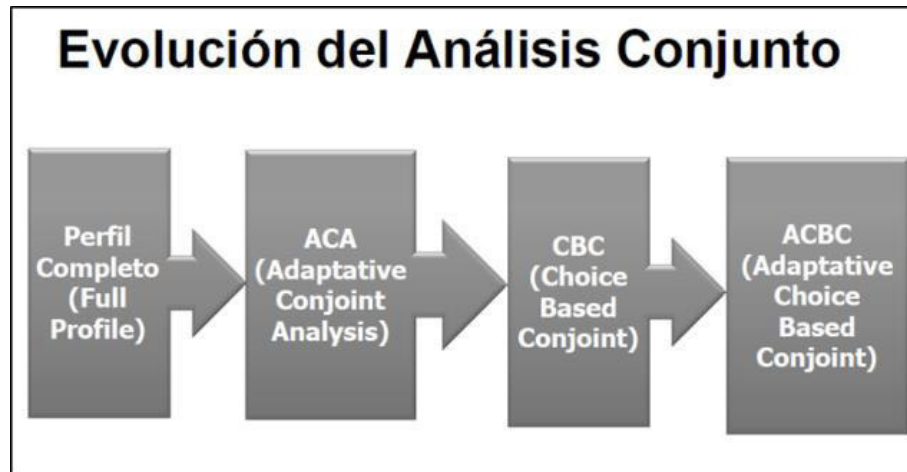
Así mismo, dentro del mismo Análisis Conjunto tradicional o clásico coexisten otras metodologías aparte del método de Perfiles Completos (Green y Rao, 1971): el Análisis Conjunto Híbrido (Green, Goldberg y Montemayor, 1981; Green, 1984), el Análisis Conjunto Adaptativo o también conocido por sus iniciales como ACA (Johnson, 1987), las Matrices de Comparaciones (Johnson, 1974) y el método Relacionado de Comparaciones Pareadas (Johnson, 1975).

Actualmente, en función a la continua aplicación del Análisis Conjunto y el surgimiento de programas informáticos que facilitan el análisis de la metodología Conjunta, se presentan nuevas metodologías: por ejemplo, el ya mencionado Análisis Conjunto Adaptativo (ACA) y el Análisis Conjunto Adaptativo Basado en la Elección. Una idea gráfica de lo expresado se observa en la figura (II.1).

El Análisis Conjunto Adaptativo (ACA) presenta mejoras en comparación al método tradicional de perfiles completos. Principalmente, a nivel del manejo de un mayor

número de atributos. De igual forma, el Análisis Conjunto Adaptativo Basado en la Elección, que es una evolución del Análisis Conjunto Basado en la Elección, se diferencia de este último por el manejo de un número mayor de atributos.

Figura II.1--Evolución del Análisis Conjunto



FUENTE: Salinas (2012)

En la figura (II.1) se identifica las metodologías actuales del Análisis Conjunto (Perfiles Completos, Análisis Conjunto Adaptativo, Análisis Conjunto Basado en la Elección y Análisis Conjunto Adaptativo Basado en la Elección). La secuencia observada describe la evolución de la técnica a lo largo de la historia.

### 2.5.1 ANÁLISIS CONJUNTO DE PERFIL COMPLETO (FULL PROFILE)

La metodología básica del Análisis Conjunto es de perfiles completos. Y su descripción recoge las siguientes características:

El Análisis Conjunto tradicional recoge preferencias (juicios) para los perfiles de productos hipotéticos, cada uno descritos en todo el conjunto de atributos seleccionados para el estudio conjunto (por ejemplo, Green y



Wind, 1975). Estos perfiles se llaman perfiles completos. Sin embargo, cuando uno concatena los niveles de todos los atributos, el conjunto de perfiles completos (o diseño factorial completo) en general será muy grande. Un encuestado estará sobrecargado indebidamente cuando se le pregunte a dar juicios de preferencia en todos los perfiles. Normalmente, se utilizan un conjunto más reducido de perfiles completos (seleccionado según un diseño experimental) en un estudio conjunto. Las preferencias declaradas globales de un individuo se descomponen en valores de utilidades independientes y compatibles que corresponden a cada atributo típicamente utilizando métodos basados en regresión. Estas funciones separadas se llaman funciones de atributo específicos partworth. En la mayoría de los casos, las funciones de preferencia pueden ser estimadas a nivel individual. Esta función de preferencia estimada puede considerarse como una función de utilidad indirecta. (Rao, 2014, p. 6)

Posteriormente, vamos a profundizar en detallar todos los aspectos de la metodología de perfiles completos ya que es la elegida como metodología conjunta para ser aplicada en la identificación de preferencias hacia un producto, en el presente estudio.

### **2.5.2 ANÁLISIS CONJUNTO ADAPTATIVO (ACA)**

Se trata de un programa informático desarrollado por Johnson en 1987. Cuya metodología es una alternativa al de los Perfiles Completos.

En 1985, Sawtooth Software lanzó su primer sistema de software de Análisis Conjunto llamado ACA (Análisis Conjunto Adaptativo). ACA se convirtió en el más popular software aplicativo de Análisis Conjunto y el método más usado tanto en Europa como en los EE.UU durante los años 1990. La principal ventaja de ACA fue su capacidad de medir más atributos de los que era recomendable con el anterior enfoque conjunto tipo tarjeta. Con ACA, fue posible estudiar una docena de dos docenas de atributos, mientras se mantengan al sujeto que indica su respuesta comprometido a proporcionar buenos datos. ACA logra esto ya que presenta diferentes secciones en la entrevista los cuales se adaptaron de las respuestas anteriores de los encuestados. En cada sección, sólo uno o pocos atributos se presentaron al mismo tiempo, para no abrumar al encuestado con demasiada información a la vez. El software lleva al encuestado a través de una investigación sistemática sobre todos los atributos, resultando en un conjunto completo de las puntuaciones de preferencia para los niveles de interés (utilidades parciales) al final de la entrevista. (Orme, 2003, p. 2)

Las ventajas y desventajas que representa la aplicación del Análisis Conjunto Adaptativo (ACA) se traducen a nivel de sus costos de implementación (relativamente altos) debido a la alta complejidad de su diseño y análisis. Además, requiere un sofisticado apoyo computacional (Cortés, 2006). En contrapartida, el ACA presenta una mayor capacidad para manejar un mayor número de atributos en comparación al procedimiento de Perfiles Completos e incluso al de la Elección Discreta (Picón et al., 2006).

### 2.5.3 ANÁLISIS CONJUNTO BASADO EN LA ELECCIÓN

El Análisis de Elección Experimental o más conocido como Análisis Conjunto Basado en la Elección (CBC, Choice-Based Conjoint por sus siglas en inglés). En los últimos años, es la metodología conjunta de mayor aceptación en ámbitos aplicados.

La metodología Basada en la Elección surge con el propósito de buscar un mayor realismo en la elección de preferencias. Debido a que en el análisis conjunto tradicional se asume que el juicio basado en las clasificaciones o calificaciones, por parte de los encuestados, recoge sus preferencias (Hair et al., 1999). Es aquí donde algunos investigadores señalan la desventaja del Análisis Conjunto tradicional respecto a esta nueva metodología, puesto que no es el modo más realista de representar el proceso de elección.

Así mismo, “el análisis conjunto basado en la elección ofrece una opción de no elección en el conjunto de elección. [...] la aproximación basada en la elección permite una contracción del mercado si todas las alternativas del conjunto de elección no son atractivas” (Hair et al., 1999, p. 439).

En la figura (II.9), se observa un ejemplo del proceso de elección haciendo uso del Análisis Conjunto Basado en la Elección. Se solicita escoger un perfil de cuatro opciones. Una de la opciones representa la opción de **no elección** describiendo verbalmente la frase “No compraría ninguna de las alternativas ofrecidas”.

Las ventajas y desventajas que representa la aplicación del Análisis Conjunto Basado en la Elección se resumen en lo siguiente: altos costos de implementación (requieren

muestras grandes para alcanzar validez estadística). Una mayor complejidad del diseño y su análisis. Y un limitado número de atributos para su manejo (menos de 6) (Cortés, 2006).

#### **2.5.4 ANÁLISIS CONJUNTO BASADO EN LA ELECCIÓN DE PERFILES PARCIALES O PARTIAL-PROFILE CBC**

El Análisis Conjunto Basado en la Elección presenta a su vez, variantes como el Análisis Conjunto Basado en la Elección de Perfiles Parciales o Partial-Profile CBC. La aplicación de esta variante representa una ventaja con respecto al Análisis Conjunto Basado en la Elección, ya que permite trabajar con muchos más atributos.

Con el Partial-Profile CBC, cada pregunta de elección incluye un subconjunto del número total de atributos estudiados. Estos atributos rotan al azar en las tareas (conjuntos de elección). Por lo que en todas las tareas de la encuesta, cada encuestado considera normalmente, todos los atributos y niveles. (Orme, 2003, p. 4)

La desventaja del Partial-Profile CBC radica en que cada tarea tiene muchas omisiones de atributos. Conllevando a que la respuesta tenga menos información. Por lo tanto, el “partial-profile CBC requiere tamaños de muestra muy grandes para estabilizar los resultados, y la estimación a nivel individual bajo el HB (Estimación Jerárquica Bayesiana) no siempre produce utilidades parciales estables a nivel individual” (Orme, 2003, p. 5).

### **2.5.5 ANÁLISIS CONJUNTO BASADO EN LA ELECCIÓN ADAPTATIVO O ADAPTATIVE CHOICE-BASED CONJOINT (ACBC)**

Esta metodología representa una evolución del Análisis Conjunto Basado en la Elección. Y surge ante la necesidad de corregir problemas suscitados a nivel de la aplicación del CBC. Al respecto, Orme (2003) señala que “si bien los métodos conjuntos basados en la elección se han popularizado en la industria. Los cuestionarios CBC pueden ser tediosos, repetitivos y sin falta de relevancia. Estos problemas, generan en muchos de los casos, que los encuestados no vean reflejados sus preferencias” (p. 5).

Sawtooth Software desarrolló un nuevo enfoque llamado CBC adaptativo. El CBC adaptativo aprovecha aspectos del ACA y el CBC. En primer lugar, se pregunta a los encuestados para que identifiquen el producto más cercano a su ideal usando un configurador (parte del principio Built Your Own - BYO “Constrúyalo Ud. Mismo”) [...] Luego, se construyen un par de docenas de conceptos de productos para que el encuestado los evalúe, todos muy similares al producto BYO. Los encuestados indican cuál de estos consideraría. Los productos considerados, se llevan a una siguiente etapa de elección para identificar al mejor, esta etapa se asemeja mucho a las visualizaciones de un CBC estándar. (Orme, 2003, p. 5)

## 2.6 LA METODOLOGIA CONJOINT

La técnica Análisis Conjunto puede ser considerada también como una metodología. Esto es, considerar al Análisis Conjunto como un método de investigación que presenta una grupo de etapas (Pérez, 2011).

Para el presente trabajo, la serie de etapas a considerar serán las proporcionadas por Varela, Rial, y Picón (2014). Estos investigadores complementan las fases, observada en la tabla (II.2), propuestas por Green y Srinivasan (1978, 1990).

Así, las etapas consideradas son:

- Planteamiento del problema.
- Selección de atributos y niveles.
- Escogiendo la metodología conjunta.
- Forma de presentación.
- Creación de estímulos.
- Trabajo de campo y tratamiento de datos.
- Estimación de las utilidades.
- Evaluación del ajuste.
- Interpretación de los resultados.

En adelante vamos a explicar al detalle cada una de las etapas antes mencionadas. Si bien estas etapas se deben cumplir en cualquier metodología conjunta, en el presente estudio las descripciones se enfocarán a la metodología tradicional de perfiles completos. El cual es la metodología elegida para llevar a cabo la aplicación de identificación de preferencias.

### **2.6.1 PLANTEAMIENTO DEL PROBLEMA**

El planteamiento del problema es considerado la primera etapa de la metodología Conjoint. Así, como su mismo nombre lo señala, en esta fase se evaluará si el problema planteado en la investigación inicial puede ser solucionado en base a la aplicación del Análisis Conjoint. Es decir, si los objetivos de la investigación se ajustan a los requeridos por un Análisis Conjoint.

Al respecto, Hair et al. (1999) señalan que el Análisis Conjoint básicamente se aplica en función a las siguientes dos finalidades: “determinar la contribución de cada atributo y sus niveles en las preferencias globales del consumidor [...], y establecer un modelo válido de los juicios de preferencia del consumidor” (p. 419).

Así mismo, Picón, Varela, y Braña (2006) expresan que el Análisis Conjoint también permite explicar los siguientes puntos, que se desprenden de las dos finalidades básicas antes mencionadas:

- Definir el producto y servicio con la mejor combinación posible de niveles de atributos.
- Identificar oportunidades de mercado mediante la exploración del potencial de productos o servicios no existentes actualmente.
- Detectar el precio óptimo para un producto concreto real o hipotético.
- Predecir las cuotas de mercado de diferentes productos con diferentes características.
- Detectar grupos de consumidores diferenciados respecto a la estructura de sus preferencias.

Algo muy importante a señalar, y que también debe ser considerado antes de aplicar el AC, es la evaluación de cómo se ha dado el proceso de decisión de los consumidores. En razón de que, el Análisis Conjunto asume, en todas sus metodologías, que el proceso de decisión de los consumidores a varias alternativas es de tipo compensatorio. Esto significa que, “un atributo desfavorable en un producto puede ser compensado con otro atributo evaluado positivamente” (Picón et al., 2006, p. 14).

Un problema que se puede presentar es que no siempre los individuos deciden de manera compensatoria. Si al analizar se verificaran estos casos, resulta conveniente decidirse por el uso de otros métodos de análisis.

### **2.6.2 SELECCIÓN DE ATRIBUTOS Y NIVELES**

La selección de atributos (también llamados factores) y sus respectivos niveles constituyen la “materia prima” del Análisis Conjunto. Por lo tanto, esta etapa constituye una fase de suma importancia.

En primer lugar, debemos enfocarnos en responder la siguiente interrogante **¿Qué atributos seleccionar?**

Al respecto, Juez y Díez (1996) nos indican que, no podemos incluir en el análisis todos los atributos que el usuario pondera a la hora de determinar la elección de un producto o aspecto frente a otro. [...] Una cuestión importante es que el objetivo no debe ser el encontrar los atributos más importantes sino aquellos que condicionan la elección. En



razón de que en muchos tipos de productos, un cierto atributo resulta tan importante e invariable que todos los productos lo ofrecen, por lo que no resulta determinante en el proceso de selección de la compra final. (p. 241)

A su vez, Picón, Varela, y Braña (2006) señalan “un atributo determinante no es solo un atributo importante para explicar las preferencias de los consumidores, sino un atributo importante que además que diferencia bien entre los producto/servicios” (p. 15).

En conclusión, no basta con seleccionar un atributo importante, sino que este debe ser un atributo que diferencie al producto o servicio del otro. Por ejemplo, al adquirir un nuevo equipo celular de una empresa de telefonía X, el comprador se centrará en evaluar atributos como: el color del equipo, marca y modelo, tamaño, etc. Y no en verificar atributos como: servicio de mensajería, si utiliza chip o no, etc. Si bien estos últimos son importantes, se entiende que todos los equipos celulares ya presentan estas cualidades.

Figura II.2-Atributos de un producto



FUENTE: Salinas (2012)

Luego, debemos saber **¿cómo identificar y seleccionar los atributos determinantes?**

Para satisfacer este objetivo existen diferentes criterios, desde “el análisis de componentes principales hasta diversos tipos de metodologías cualitativas como las entrevistas en profundidad, los juicios de expertos (técnica Delphi, etc.), o los grupos de discusión con consumidores” (Picón et al., 2006, p. 15).

Una vez seleccionados los atributos determinantes del producto o servicio nos queda la tarea de elegir los niveles que se considerarán en cada uno de los atributos. Uno de los primeros aspectos es el de considerar el tipo de atributo. Ya que “no es lo mismo un atributo categórico [...] en el que los niveles posibles son limitados y claros en la mayoría de las ocasiones que un atributo continuo (por ejemplo el precio) en el que el rango de elección de niveles es infinito” (Juez & Díez, 1996, p. 242). Así mismo, el número de niveles a seleccionar condiciona el proceso de análisis posterior, por ejemplo el número de parámetros a estimar o el número de perfiles (estímulos) que se presentan al encuestado. Al respecto, Picón, Varela, y Braña (2006) indican también que “cuando todos los atributos tienen los mismos niveles, los valores absolutos de las utilidades parciales ofrecen una indicación de la importancia relativa de los atributos [...]. Un atributo adquiere más importancia a medida que aumenta el número de sus niveles, incluso aunque los niveles mínimo y máximo sean los mismos” (p. 16).

Bryan Orme (2002) en su investigación “Formulating attributes and levels in conjoint analysis” propone tener en cuenta una serie de consideraciones al momento de generar atributos y niveles:

1. Los niveles deben ser declaraciones concisas con significado concreto.
2. Los atributos deben ser independientes.

3. Los niveles dentro de cada atributo deben ser mutuamente excluyentes.
4. Los niveles de los atributos deben cubrir toda la gama de posibilidades para los productos existentes así como los productos que aún no existan, pero que desea investigar.
5. Las prohibiciones deben usarse con moderación, o no.
6. El número de niveles que escoja para definir un atributo puede tener una influencia significativa en los resultados.
7. Los atributos que no pueden ser adecuadamente descritos en palabras deberían estar representados en multimedia.
8. En nuestra experiencia, los encuestados tienen dificultades para hacer frente a más de seis atributos en los métodos conjuntos de perfiles completos (CVA) así como el CBC.

Figura II.3-Selección de atributos y niveles: ejemplo viaje de promoción

Atributos	Niveles de atributos
Destino	<ol style="list-style-type: none"> <li>1. Varadero</li> <li>2. Punta Cana</li> <li>3. Cancún</li> <li>4. San Andrés</li> </ol>
Precio	<ol style="list-style-type: none"> <li>1. \$ 600</li> <li>2. \$ 700</li> <li>3. \$ 800</li> </ol>
Alojamiento	<ol style="list-style-type: none"> <li>1. Hotel 4 estrellas, todo incluido</li> <li>2. Hotel 4 estrellas, sólo desayuno</li> <li>3. Hotel 3 estrellas, sólo desayuno</li> </ol>
Forma de pago	<ol style="list-style-type: none"> <li>1. 20% de descuento 3 meses antes</li> <li>2. 10% de descuento 2 meses antes</li> <li>3. 5% de descuento 1 mes antes</li> <li>4. Sin descuento por pago en los últimos 30 días</li> </ol>
Duración del viaje	<ol style="list-style-type: none"> <li>1. 6 días</li> <li>2. 7 días</li> <li>3. 8 días</li> </ol>

FUENTE: Salinas (2012)

### 2.6.3 ESCOGIENDO LA METODOLOGÍA CONJUNTA

Esta etapa consiste en escoger la metodología del Análisis Conjunto que se va a utilizar. Como ya se ha visto, en la actualidad coexisten dos criterios básicos: el Análisis Conjunto tradicional y Análisis Conjunto Basado en la Elección (Picón et al., 2006).

Así mismo, se observó en la figura (II.1) que nuevas metodologías se han ido desarrollando como una evolución de los variantes básicas. En razón, principalmente, del desarrollo de programas informáticos que han automatizado los procesos conjuntos. Permitiendo un manejo de un mayor número de atributos. Así tenemos al Análisis Conjunto Adaptativo (ACA) y el Análisis Conjunto Basado en la Elección Adaptativo o también conocido como Adaptive CBC (ACBC).

Pero en qué se fundamenta el criterio para escoger una metodología respecto de la otra. Varela, Rial, y Picón (2014) señalan que se optará por una u otra metodología dependiendo, principalmente, de la respuesta obtenida a la pregunta “clave”

**¿CUÁNTOS ATRIBUTOS VAN A INCLUIRSE EN EL ESTUDIO?.**

Al respecto, Hair et al. (1999) sostienen que el “Análisis Conjunto tradicional [...] se caracteriza por ser un modelo aditivo simple que contiene 9 factores estimados para cada individuo. [...] El método adaptativo conjunto se desarrolló para dar cabida a un gran número de factores (muchas veces más de 30) que no sería factible en un Análisis Conjunto tradicional” (p.420).

Tabla II.4-Comparación de metodologías conjuntas alternativas

Características	Metodología conjunta		
	Tradicional	Adaptativo	Basado en elección
<b>Número máximo de atributos</b>	9	30	6
<b>Nivel de análisis</b>	Individual	Individual	Agregado
<b>Forma del modelo</b>	Aditivo	Aditivo	Aditivo + efectos interacción

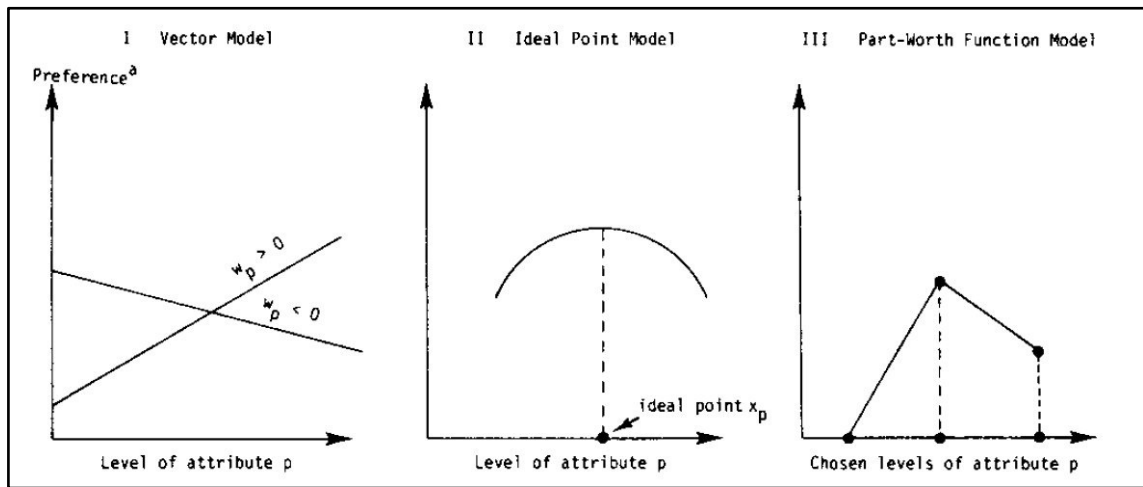
FUENTE: Hair et al. (1999)

Por su parte, los investigadores Varela et al. (2014) recomiendan utilizar el Análisis Conjunto Adaptativo (ACA) cuando se considere 10 atributos o más. Los Perfiles Completos cuando se presente menos de 10 atributos. Y, el Análisis Conjunto Basado en la Elección (CBC), cuando se presente 6 atributos o menos.

La elección de la metodología del Análisis Conjunto implica, también, la selección de un modelo o función de preferencia. Ya que las preferencias no son las mismas para los distintos niveles de atributos. Al respecto, se debe entender por modelo de preferencia a "la relación funcional que liga la utilidad total de una configuración determinada de atributos con las utilidades parciales de cada uno de esos atributos y niveles" (Juez & Díez, 1996, p. 243).

Los modelos de preferencia (ver figura II.4) se clasifican según lo propuesto por Green y Srinivasan (1978). El Modelo Vectorial (que establece una relación lineal entre los atributos y las preferencias), el Modelo de Punto ideal (establece una relación de tipo cuadrático) y el Modelo de Pesos Parciales (establece una relación lineal por partes). Adicional a estos se deben considerar también los Modelos Mixtos.

Figura II.4-Modelos de preferencia



FUENTE: Green and Srinivasan (1978)

Según Picón et al. (2006) la selección del modelo de preferencia debe evaluar lo siguiente:

- Se debe escoger el modelo que presente una mayor validez predictiva. Es decir, el de menor error de predicción.
- La solución pasa primero por echar un vistazo a la forma de la función que asocia los atributos con las preferencias.
- Cuando el atributo es categórico se debe considerar el Modelo de la Función de Pesos Parciales.
- Otra forma de seleccionar el modelo de preferencia mas adecuado (en un análisis de perfiles completos, diseño factorial fraccionado y regresión múltiple, que es uno de los más comunes) consiste en el uso de test estadísticos de comparación de modelos.

Lo normal es hacer operativo a estos modelos de preferencia de manera aditiva. Esto significa suponer que cada atributo tiene un efecto independiente de los demás en la

formación de preferencias y los valores de cada atributo se suman para obtener la utilidad total del perfil (Picón et al., 2006).

#### **2.6.4 FORMA DE PRESENTACIÓN**

Existen diferentes formas de recoger la información ya que esta varía según la metodología escogida.

Según Picón et al. (2006) hay cinco modos, básicamente de recojo de información. En primer lugar, a nivel de Análisis Conjunto “clásico” se presentan las matrices de comparaciones, también conocidas como matrices trade-off, consistente en una serie de tarjetas en cada una de las cuales aparecen combinados sólo dos atributos con sus respectivos niveles. La tarea consiste en ordenar las múltiples combinaciones de dos niveles, de más a menos preferida. Las ventajas de esta presentación consisten en su sencillez y rápida clasificación de las distintas combinaciones. Por el contrario, las desventajas son, fundamentalmente, la falta de realismo que significa el presentar sólo dos atributos a la vez y el excesivo número de juicios de preferencias exigidos al encuestado.

Un ejemplo de matrices de comparaciones lo observamos en la publicación de Bryan Orme (2006) en “A short history of conjoint analysis. Getting Started with Conjoint Analysis”. Donde se comparan dos atributos sobre automóviles: lugar de procedencia y tracción del vehículo. La tarea del encuestado consiste en asignar un orden que va desde 1 hasta el 9 según su preferencia ante la combinación observada (ver figura II.5).

Figura II.5-Matriz de comparaciones

<b>Exposición 2: Matriz de Comparaciones de Johnson</b>			
	<b>Hecho en EE.UU.</b>	<b>Hecho en Europa</b>	<b>Hecho en Japón</b>
<b>Tracción delantera</b>	<b>7</b>	<b>6</b>	<b>3</b>
<b>Tracción trasera</b>	<b>9</b>	<b>8</b>	<b>5</b>
<b>Todas las ruedas motrices</b>	<b>4</b>	<b>2</b>	<b>1</b>

FUENTE: Orme (2006)

Un segundo método, también a nivel del Análisis Conjunto tradicional, consiste en la ordenación de Perfiles Completos. En primer lugar, se elabora una serie de descripciones de un producto o servicio con todos los atributos y utilizando un solo nivel para cada atributo. La tarea consiste en ordenar los perfiles de más o menos preferidos. Las ventajas de este criterio es que los perfiles pueden ser presentados verbalmente, mediante tarjetas, diseños de ordenador, productos reales, etc.

Un ejemplo sobre este tipo de forma de presentación se observa en la investigación realizada por Schnettler et al. (2010). “Aceptación hacia yogurt con diferentes ingredientes funcionales en consumidores de supermercados del sur de Chile”. A los encuestados se les solicitó ordenar los distintos perfiles (se observa un total de ocho estímulos) desde la más preferida hasta la menos preferida usando una escala del 1 al 8. Las preferencias en el estudio se dieron de manera inversa. Ya que el 1 representó la opción más preferida y el valor 8, la menos preferida. Las tarjetas con las distintas combinaciones para ser ordenados por los encuestados se muestran en la figura (II.6).



Figura II.6-Ordenación de perfiles completos

<b>Alternativa A</b>  <b>\$110</b>	<b>Alternativa B</b>  <b>\$125</b>	<b>Alternativa C</b>  <b>\$100</b>	<b>Alternativa D</b>  <b>\$100</b>
<b>Alternativa E</b>  <b>\$100</b>	<b>Alternativa F</b>  <b>\$125</b>	<b>Alternativa G</b>  <b>\$100</b>	<b>Alternativa H</b>  <b>\$110</b>

FUENTE: Schnettler et al. (2010)

Otra forma de presentación y recolección de los datos a nivel del Análisis Conjunto tradicional es la cuantificación de perfiles completos. El análisis es muy similar al criterio anterior. Lo que si cambia es la tarea a nivel de los encuestados. En este caso particular, se les pide que expresen numéricamente el grado de preferencia por cada uno de los perfiles evaluados. (ver figura II.7)

Figura II.7-Cuantificación de perfiles completos

**¿Cuán dispuesto a comprar estaría usted un computador con las siguientes especificaciones?**

**Presario 700**  
**Pentium II 900 Mhz**  
**Monitor 14 pulgadas**  
**256 Mb RAM**  
**20 Gb Disco Duro**  
**CDR-DVDR**  
**\$ 900.000**

Definitivamente No compraría      Indiferencia      Definitivamente compraría

☐ 1   
 ☐ 2   
 ☐ 3   
 ☐ 4   
 ☐ 5   
 ☐ 6   
 ☐ 7   
 ☒ 8   
 ☐ 9

FUENTE: Cortés (2006)

A nivel del Análisis Conjunto Adaptativo, se establece un modo de recojo denominado Comparaciones Pareadas. Este criterio es un método a intermedio entre los perfiles completos y las matrices trade-off. Consiste en comparar dos perfiles incompletos formadas por un subconjunto del total de atributos evaluados en el estudio. Al igual que en las matrices de comparaciones los encuestados evalúan estímulos de dos en dos, pero aquí los estímulos son perfiles en vez de atributos. Su evaluación se da sólo mediante un ordenador (Varela et al., 2014). Un ejemplo se observa en la figura (II.8).

Figura II.8-Presentación de estímulos en ACA

¿Cual de los siguientes vehículos usted estaría mas dispuesto a comprar?

 <p><b>Toyota Rav4</b> \$ 23,000 2300 cc</p>		 <p><b>Kia Sportage</b> \$ 21,000 1800 cc</p>						
Definitivamente prefiero A		Indiferencia		Definitivamente prefiero B				
1	2	3	4	5	6	7	8	9

FUENTE: Salinas (2012)

Por último, Picón et al. (2006) señalan sobre los conjuntos de elección que, característicos de la aproximación Basada en la Elección, “aquí los sujetos no necesitan ordenar ni mucho menos cuantificar todas las alternativas disponibles. Simplemente han de escoger el perfil más preferido entre un pequeño subconjunto del total, pudiendo además no elegir ninguna de la opciones presentadas” (p. 33). De forma similar al Análisis Conjunto Adaptativo, el proceso de recojo de los datos se da sólo a nivel de un ordenador.

Figura II.9-Presentación de estímulos en un Análisis Conjunto Basado en la Elección

**¿Cuál de los siguientes computadores usted estaría más dispuesto a comprar?**

<b>Computador A:</b>  Presario 700 PII 400 Mhz Monitor 14' 64 Mb RAM 5 Gb Disco Duro CDRW-DVDRW \$600.000	<b>Computador B:</b>  Presario 700 PIII 900 Mhz Monitor 14' 256 Mb RAM 20 Gb Disco Duro CDR-DVDR \$900.000	<b>Computador C:</b>  Dell 5000 PIV 2 Ghz Monitor 21' 1 Gb RAM 120 Gb Disco Duro - \$1.500.000	No compraría ninguna de las alternativas ofrecidas.
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>

FUENTE: Cortés (2006)

### 2.6.5 CREACIÓN DE ESTÍMULOS

Según la metodología conjunta escogida se podrá establecer un determinado número de estímulos (combinaciones). Por ejemplo, si consideramos el método de Matrices de Comparaciones, el número de matrices a evaluar por los individuos está en función al número de atributos considerados y se incrementa considerablemente de acuerdo al aumento del número de atributos tomados en cuenta en el estudio. La forma de encontrar el número total de matrices de comparaciones se da en base a la fórmula (II.1).

Fórmula II.1-Número de matrices de comparaciones

$$\binom{j}{2} = \frac{j!}{2! (j-2)!}$$

FUENTE: Picón et al. (2006)

Donde  $j$  es el número de atributos.

Por otro lado, si consideramos ahora el método de perfiles completos la cantidad de perfiles se incrementa considerablemente de acuerdo a la cantidad de atributos y número de niveles tomados en cuenta. En conclusión, nos encontramos ante un primer problema el cual consiste en la saturación de información proporcionado a los sujetos para su evaluación.

Como ya se ha señalado, la metodología del Análisis Conjunto abordado en el presente estudio se enfoca al Análisis Conjunto tradicional de perfiles completos. Por lo tanto, se explicará de manera pormenorizado sobre la creación de estímulos bajo estos criterios.

En el Análisis Conjunto de perfiles completos el número de combinaciones posibles llamado diseño factorial completo depende del número de atributos y niveles considerados. Al respecto, Juez y Díez (1996) señalan lo siguiente: “una vez que se ha optado por elaborar tarjetas, o más genéricamente estímulos, siguiendo un esquema de perfil completo, hay que decidir que combinaciones de niveles de los atributos van a ser presentados al encuestado. Evidentemente, no pueden elaborarse tanto estímulos como combinaciones de niveles posibles” (p. 251). Donde el número de combinaciones bajo el modo de perfil completo se obtiene con la fórmula (II.2).

Fórmula II.2-Número de combinaciones en el método de perfiles completos

$$\textit{Combinaciones} = m^p . n^q \dots$$

FUENTE: Juez and Díez (1996)

Donde:

$m$ : número de niveles

$p$ : número de atributos con  $m$  niveles

$n$ : número de niveles

$q$ : número de atributos con  $n$  niveles

\*: ....  
\*: ....

Si consideramos el ejemplo del Viaje de Promoción (figura II.3) para conocer el número total de combinaciones posibles; se observa, en primer lugar, un total de 5 atributos considerados (destino, precio, alojamiento, forma de pago y duración del viaje). De los cuales 3 atributos presentan 3 niveles cada uno y 2 atributos presentan 4 niveles cada uno. Por lo tanto, el número de combinaciones obtenidas se daría del siguiente modo:  $\text{Combinaciones} = 3^3 \cdot 4^2$  resultando un total de 432 combinaciones. Estamos frente a un caso de Diseño Factorial Completo.

Los esfuerzos se redireccionarían a reducir la excesiva información brindada al individuo.

Picón et al. (2006) señalan que disponemos de otros procedimientos que nos permiten reducir el número de combinaciones que se le presentan al sujeto sin perder por ello la capacidad para estimar los parámetros:

1. Diseño factoriales fraccionados ortogonales, para reducir el número de tarjetas que se presentan al sujeto,
2. Métodos alternativos (enfoques autoexplicados, análisis conjunto híbrido, análisis conjunto adaptativo) para reducir el número de factores en cada tarjeta (p.24).

Para nuestro caso se utilizará el criterio de los diseños factoriales fraccionados ortogonales. Pero, en primer lugar definamos qué es un diseño factorial fraccionado. Según Juez y Díez (1996) señalan que “el diseño factorial fraccionado nos permite

reducir el número de estímulos a presentar al encuestado, a expensas de confundir los efectos (o parámetros de la función principal” (p. 253). A su vez, Picón et al. (2006) indican que “los investigadores suelen utilizar con frecuencia algún tipo de diseño factorial fraccionado, con un menor número de combinaciones pero manteniendo la eficiencia (bondad) a la hora de estimar los parámetros” (p. 24). Considerando principalmente que la eficiencia de un diseño está en función de las varianzas y covarianzas de las estimaciones. Y ésta se incrementa a medida que disminuyen las varianzas (Picón et al., 2006).

La tarea ahora consiste en encontrar el diseño factorial fraccionado que presente una mayor eficiencia. En ese objetivo se considera al diseño factorial fraccionado ortogonal que es un tipo especial de diseño fraccionado.

Los diseños que solemos denominar ortogonales se caracterizan de hecho por ser a la vez ortogonales y equilibrados. [...] un diseño es ortogonal cuando la ocurrencia conjunta de cualesquiera dos niveles de diferentes atributos es igual al producto de sus frecuencias marginales [...] y será equilibrado cuando cada nivel aparece dentro de cada factor un número igual de veces (Picón et al., 2006, p. 25).

Estos diseños también conocidos como Diseños Completamente Ortogonales presentan “una eficiencia óptima y se suelen utilizar para estimar de forma independiente los efectos principales del modelo aditivo conjunto (el más frecuente en el análisis conjunto tradicional) con el menor número de perfiles posibles” (Picón et al., 2006, p. 25).

Así mismo, otra consideración importante a ser tomado en cuenta cuando se presenta una cantidad reducida de estímulos es la planteada por Hair et al. (1999) quienes señalan que “el número de atributos factores en el análisis afecta directamente a la eficiencia estadística y a la fiabilidad de los resultados. A medida que se añaden más factores y niveles, el creciente número de parámetros a estimar exige o bien un número mayor de estímulos o bien una reducción de la fiabilidad de los parámetros” (pp. 421-422). Y plantean que, el número de estímulos a ser evaluados por cada encuestado, si el análisis se realiza a nivel individual, es de acuerdo a la fórmula (II.3).

Fórmula II.3-Número mínimo de estímulos

$$T_{min} = N_{niv} - N_{fac} + 1$$

FUENTE: Hair et al. (1999)

Siendo:

$T_{min}$  : Número mínimo de estímulos

$N_{niv}$  : Número total de niveles para todos los factores

$N_{fac}$  : Número total de factores

Si se considera nuevamente el ejemplo del Viaje de promoción (Figura II.3) en donde se observaba un total de 5 atributos (destino, precio, alojamiento, forma de pago y duración del viaje). De los cuales 3 atributos presentan 3 niveles cada uno y 2 atributos presentan 4 niveles cada uno. Y se obtenía un total de 432 combinaciones. Aplicando la fórmula (II.3) del número mínimo de estímulos se obtendría  $(17-5+1)$  estímulos. Esto es, tendríamos 13 estímulos (combinaciones) como mínimo para ser evaluado por cada individuo.

### 2.6.6 TRABAJO DE CAMPO Y TRATAMIENTO DE DATOS

Si bien ya se había mencionado las diversas formas de recogidas de datos, según las diversas metodologías aplicadas, en el subcapítulo Forma de presentación. En este subtema, se va a abordar la **forma de evaluar las preferencias** sobre las combinaciones en la metodología de perfiles completos.

Existen dos procedimientos principales: la ordenación o ranking de los perfiles completos, y la cuantificación o puntuación de los perfiles completos.

- **Ordenación o ranking de los perfiles completos**, bajo este criterio el sujeto debe ordenar los estímulos de mayor a menor preferencia o viceversa. Así mismo, se considera un ranking que va desde 1 hasta el número total de combinaciones establecidas. Por lo tanto, la función de los encuestados es establecer ese orden de preferencias sobre el conjunto de combinaciones.
- **Cuantificación o puntuación de los perfiles completos**, muy similar al anterior criterio, lo que cambia es la función del encuestado. Ya que tendrá que asignar un puntaje que representa el grado de preferencia por cada uno de las combinaciones. Al respecto, no hay un criterio único de establecer la escala de puntuaciones, así podemos observar escalas que van de 1 a 5, 1 a 10, 1 a 100, etc.

A nivel del trabajo de campo, el Análisis Conjunto de perfiles completos presenta muchas ventajas con respecto a las demás metodologías. Uno de ellos se observa a nivel de la recogida de datos, ya que no necesita de la computadora para cumplir esta función. Como sí es el caso del ACA y del Análisis Conjunto Basado en la Elección.



Figura II.10-Ejemplo de ordenación de perfiles completos



FUENTE: Varela et al. (2014)

### 2.6.7 ESTIMACIÓN DE LAS UTILIDADES

Existen diversos criterios para la estimación de las utilidades. Estos dependen de la escala de medición de la variable dependiente.

Según Green y Srinivasan (1978) estas se clasifican en tres categorías: cuando la variable dependiente está en escala ordinal, los métodos en esta clase son MONANOVA (Kruskal, 1965), PREFMAP (Carroll, 1972), procedimiento de Johnson para matrices de comparaciones (Johnson, 1973; Nehls, Seaman, and Montgomery 1976), y LINMAP (Srinivasan y Shocker 1973a, 1973b; Pekelman y Sen 1974). Cuando la variable dependiente está en escala de intervalo se utiliza más comúnmente el método de Mínimos Cuadrados Ordinarios (OLS). Y si se recogió los datos por mediante la elección de la alternativa más preferida, los métodos más adecuados son LOGIT (McFadden, 1976; Ben-Akiva, 1973; Gensch, Golob, and Recker 1976; Green and Cannone, 1977; Punj and Staelin, 1978) y PRO BIT (Goldberger 1964, pp. 250-1; Rao and Winter 1977).

De los métodos antes mencionados cada uno de ellos presentan ventajas y desventajas. A nivel de los métodos métricos Picón et al. (2006) señalan que “el procedimiento OLS tiene la ventaja de proporcionar errores estándar para los parámetros estimados lo que no hace de modo directo ninguno de los métodos anteriores” (p. 37). Así mismo, mencionan que “diversos estudios han encontrado que la regresión OLS aplicada a datos no métricos (con variables ficticias) producen resultados similares en cuanto a validez predictiva a los obtenidos mediante algoritmos no métricos, más complejos y difíciles de encontrar” (p.37). Los investigadores Darmon y Rouzies (1991) verificaron que el método de Mínimo Cuadrados Ordinarios (OLS) proporcionaba las estimaciones de importancia más robustas, en comparación a los métodos LINMAP o MONANOVA. Por esta razón, “la regresión OLS ha ido imponiéndose como el método de estimación conjunto por excelencia para el Análisis Conjunto Tradicional” (Picón et al., 2006, p. 38).

Dentro del proceso de estimación se busca encontrar los “part-worth” o también conocidos como “utilidades parciales”. Los cuales relacionan los distintos niveles de los atributos considerados con las preferencias de los encuestados bajo un modelo aditivo lineal el cual es el más utilizado en el Análisis Conjunto.

En este sentido, Cortés (2006) indica que “el total de utilidad de un perfil en un modelo de preferencias compensatorio es igual a la suma de los part-worth” (sección de Construcción del modelo, párr. 2).

Fórmula II.4-Modelo aditivo lineal

$$U_i = \sum_{j=1}^J u_{jk}$$

FUENTE: Picón et al. (2006)

Donde

$U_i$  es la utilidad total que muestra un individuo por el perfil  $i$

$u_{jk}$  es la utilidad parcial asociada con el nivel concreto  $k$  del atributo  $j$

Con respecto a los modelos aditivos, Rao (2014) indica que el modelo aditivo de utilidad es aplicable para un grupo de datos que presenta un arreglo ortogonal o cuando las interacciones entre los atributos son ignorados. Este modelo de utilidad estará compuesto correspondientemente por tres tipos de atributos. Dejando el número de atributos de estos tres tipos (categórico, vector y punto ideal) ser  $(p_1+p_2+p_3=p)$ , la función de utilidad puede ser especificada como:

$U=\beta_0+$  funciones de las utilidades parciales para los  $p_1$  atributos que son categóricos + funciones de las utilidades parciales para los  $p_2$  atributos que son del tipo vector + funciones de las utilidades parciales para los  $p_3$  atributos que son del tipo punto ideal. (p.82)

El modelo básico de regresión presenta la siguiente forma general (fórmula II.5).

Fórmula II.5-Modelo básico de regresión

$$y_i = \beta_0 + \sum_{j=1}^J \beta_j x_{ij} + e_i$$

FUENTE: Picón et al. (2006)

El modelo presentado en la fórmula (II.5) puede representarse, a su vez, según la forma presentada en la fórmula (II.6).

Fórmula II.6-Modelo de preferencias globales

$$\gamma_i = \beta_0 + \sum_{j=1}^J \hat{u}_{jk}$$

FUENTE: Picón et al. (2006)

Donde

$\gamma_i = \hat{U}_i$  es la respuesta (observación de preferencia) de un sujeto para el perfil  $i$

$\beta_0$  es la constante de regresión

La especificación y estimación de las utilidades parciales  $u_{jk} = (\beta_j x_{ij})$  van a variar de acuerdo al modelo que siguen las preferencias de los individuos por los diferentes niveles de un atributo (discreto, lineal, ideal o anti-ideal) (Picón et al., 2006).

Al respecto, Pérez (2011) señala que, en cuanto a las columnas que representan los atributos o factores, cabe matizar que si, por ejemplo,

para ese atributo se definieron  $m_i$  niveles discretos, entonces dicho atributo o factor dará lugar a  $m_i-1$  columnas. Estas columnas serán utilizadas para estimar los  $m_i-1$  valores de  $\alpha_{ij}$ . Por el contrario, si los niveles de ese factor o atributo son lineales, entonces habrá una columna de valores centrados de ese factor  $(x_{ij} - \bar{X}_i)$ . Dichas columnas serán usadas para estimar los valores  $\hat{\beta}_i$  (p. 453).

Además, Varela et al. (2003) indican que para la estimación de utilidades deben tomarse en cuenta las siguientes consideraciones:

a) Si los factores o atributos son considerados discretos (categóricos) se establecería una relación de **Componentes parciales separados** (ver figura II.11). Donde, se recurre a una regresión con variables ficticias o dummy, asignando valores arbitrarios a los diferentes predictores (los niveles de atributo) que sirvan para diferenciarlos. Un ejemplo de factor discreto lo tenemos en el Entorno que define un destino turístico: Playa, Naturaleza, Montaña y Ciudad (Picón et al., 2006), entonces:

$$\hat{u}_{jk} = \begin{cases} \hat{a}_{jk} & \text{para todo } k = 1, \dots, m_j - 1 \\ \sum_{j=1}^{m_j-1} \hat{a}_{jk} & \text{para todo } k = m_j \end{cases}$$

Donde:

$\hat{u}_{jk}$  son la utilidades parciales o partworth

$\hat{a}_{jk}$  son los  $j$  niveles de los  $k$  factores discretos

b) Si los factores son considerados lineales (ver figura II.11) se recurre a una regresión lineal estándar y solamente será necesario estimar un parámetro: la pendiente de la recta  $\beta_j$ . Un ejemplo de factor considerado lineal lo observamos en el atributo Garantía, ya que en principio siempre ocurre que a mayor garantía de un producto, mayor preferencia del mismo. (Picón et al., 2006). Luego,

$$\hat{u}_{jk} = \hat{\beta}_j X_k$$

Donde:

$X_k$  son los factores lineales

c) Si por el contrario, asumimos que entre los factores y atributos definidos y las preferencias de los sujetos siguen un modelo ideal o anti-ideal (ver figura II.11), entonces:

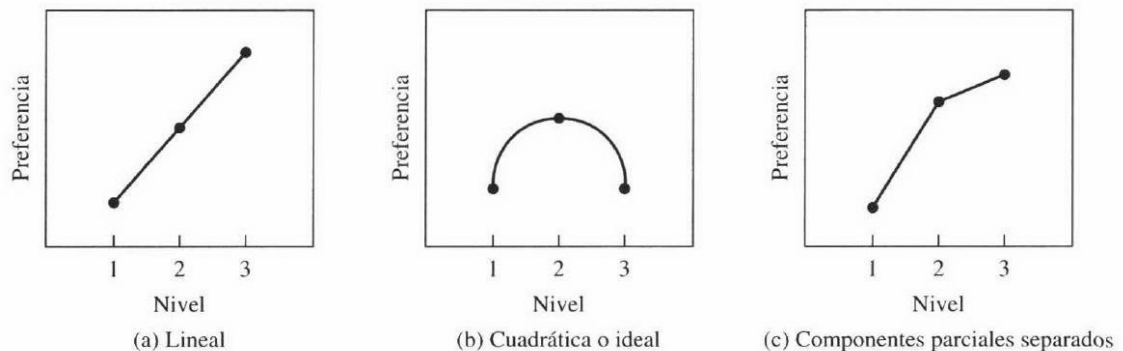
$$\hat{u}_{jk} = \hat{\gamma}_{j1} Z_{jk} + \hat{\gamma}_{j2} Z_{jk}^2$$

Donde:

$Z_{jk}$  son los  $j$  factores ideal o anti-ideal

Un ejemplo de factor que sigue un modelo de punto ideal lo observamos a nivel de las preferencias de los productos dulces, siendo rechazados los muy como los poco azucarados (Picón et al., 2006)

Figura II.11- Los tres tipos de relaciones básicas entre los niveles de los factores en el AC



FUENTE: Hair et al. (1999)

Por lo tanto, se podrán modelar las respuestas de preferencia de cada encuestado como funciones lineales de  $\beta$ . Algo importante a señalar es que se van a obtener una determinada cantidad de ecuaciones tanto como el número de perfiles se presenten al individuo.

Fórmula II.7-Modelo de preferencias en su forma matricial

$$\tau = X\beta + e$$

FUENTE: Picón et al. (2006)

Donde  $\tau' = [\tau_1, \tau_2, \tau_3, \tau_4, \dots, \tau_j]$  es el vector de respuestas  $\tau_i$ ;  $\beta' = [\beta_0, \beta_1, \beta_2, \beta_3, \dots, \beta_j]$  es el vector de parámetros a estimar; X es una matriz dada por:

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1j} \\ 1 & x_{21} & x_{22} & \dots & x_{2j} \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{i1} & x_{i2} & \dots & x_{ij} \end{bmatrix}$$

y  $e' = [e_1, e_2, e_3, e_4, \dots, e_j]$  representa el vector de residuos. Donde la solución para  $\hat{\beta}$  es única y directa:  $\hat{\beta} = (X'X)^{-1}X'\tau$  (Picón et al., 2006, p. 40).

Por último, las utilidades parciales estimadas permiten conocer la importancia relativa, el cual permite comparar la importancia de cada atributo considerado en el estudio. Esta importancia se obtiene a través de la fórmula (II.8). Pero antes de calcular la importancia relativa se debe calcular la importancia de cada atributo.

La importancia de un atributo es el valor absoluto de la diferencia entre la utilidad más alta y la utilidad más baja de los niveles de los atributos:  $IMP_i = |\max(\beta_{ij}) - \min(\beta_{ij})|$

Donde:  $\beta_{ij}$  es el valor de la utilidad asociado al nivel j del atributo i.

Fórmula II.8-Importancia relativa de un atributo

$$IMP_j = 100 \frac{RANGO_j}{\sum_{j=1}^J RANGO_j}$$

FUENTE: Picón et al. (2006)

### 2.6.8 EVALUACIÓN DEL AJUSTE

En este subcapítulo vamos a explicar lo relativo a la fiabilidad y validez del estudio conjunto. Al respecto Picón et al. (2006) señalan que se pueden plantear varias medidas de fiabilidad en un estudio conjunto:

1. **La fiabilidad temporal**, donde se trata de responder a la pregunta ¿se hubiesen obtenido los mismos resultados en otro momento temporal distinto?.
2. **La fiabilidad del conjunto de estímulos**, donde se trata de responder a la pregunta ¿se hubiesen obtenido los mismos resultados utilizando un conjunto diferente de estímulos o perfiles?
3. **La fiabilidad del conjunto de atributos**, donde se trata de responder a la pregunta ¿las utilidades para un conjunto de atributos hubiesen sido las mismas si esos atributos hubiesen estado acompañados de otros atributos diferentes?
4. **La fiabilidad del método de recogida de datos pregunta**, donde se trata de responder a la pregunta ¿se hubiesen obtenidos los mismos resultados de haberse utilizado otros procedimientos de recogida de datos?

El objetivo de cuantificar la fiabilidad es el de medir el grado de precisión de las utilidades estimadas, Para este fin, se han propuesto diversas medidas de fiabilidad con sus ventajas y desventajas, “pero la más utilizada hasta ahora sigue siendo la correlación producto-momento de Pearson entre las utilidades (normalizadas) obtenidas de la



evaluación principal y las obtenidas de la evaluación utilizada para testar la fiabilidad” (Picón et al., 2006, p. 48).

Luego, a nivel de la validez del análisis conjunto, considerando que este tipo de análisis desarrolla un modelo con fines de predicción de preferencias, se pueden medir de las siguientes maneras:

- **Validación interna**, (la medición se obtiene a través del coeficiente de determinación  $R^2$  de Pearson o la Tau de Kendall). A nivel de este tipo de validación se busca evaluar el ajuste general del modelo. Es decir, el ajuste entre las utilidades obtenidas por el AC y el orden o puntuación inicial establecida por cada encuestado. Además, “la validación interna implica la confirmación de que la regla de composición seleccionada (es decir, aditiva frente a interactiva) es la apropiada” (Hair et al., 1999, p. 436).
- **Validación cruzada**, (la medición se obtiene a través de los estímulos holdout). Estos estímulos o perfiles holdout son un conjunto de perfiles de reserva o de validación. “La validación cruzada es el procedimiento más utilizado para evaluar la capacidad predictiva del Análisis Conjunto” (Picón et al., 2006, p. 49).
- **Validación externa**, al respecto Hair et al. (1999) indican que “esta validación implica en general la capacidad del análisis conjunto de predecir elecciones efectivas, y de forma más específica la representatividad de la muestra” (p.436).

Así mismo, Picón et al. (2006) señalan que para el análisis de esta “auténtica” forma de validez predictiva se han desarrollado tres enfoques diferentes que se mencionan a continuación:

- Comparaciones entre la cuota del mercado predicha por un simulador de elección y las cuotas de mercados actuales o futuras.
- Comparaciones a nivel individual entre los resultados de los análisis y la intención o el comportamiento de compra del sujeto, y
- Comparaciones a nivel individual entre los resultados de los análisis y las selecciones posteriores del sujeto respecto a los productos evaluados.

Con respecto a este último tipo de validación (externa), muy pocos estudios han desarrollado una evaluación considerando estos procedimientos.

# **CAPÍTULO III**

## **APLICACIÓN Y RESULTADOS**

### **3.1 INTRODUCCIÓN**

En el presente capítulo vamos a desarrollar una aplicación del Análisis Conjunto tradicional de perfiles completos. A través de este caso práctico se busca conocer las preferencias que tienen los consumidores hacia un producto: Identificando las preferencias académicas universitarias de los alumnos de los últimos años de educación secundaria.

Algo importante a señalar es que todo el proceso se llevará a cabo en función a las etapas de la metodología conjunta, mencionadas en el capítulo II.

- Planteamiento del problema.
- Selección de atributos y niveles.
- Escogiendo la metodología conjunta.
- Forma de presentación.
- Creación de estímulos.
- Trabajo de campo y tratamiento de datos.

- Estimación de las utilidades.
- Evaluación del ajuste.
- Interpretación de los resultados

Así mismo, antes de comenzar las fases de esta aplicación cabe mencionar que el análisis estadístico de todo el proceso se llevará acabo utilizando el programa estadístico R en su versión 3.1.1. Por lo tanto, a nivel de este capítulo se mencionarán y explicarán los comandos y resultados obtenidos realizados por este aplicativo.

### **3.2 PLANTEAMIENTO DEL PROBLEMA**

El tema de investigación es la identificación de preferencias universitarias. Esto, hay que entenderlo desde un punto de vista referido a la elección de la universidad de preferencia y no a hacia la carrera o algún curso universitario.

La población del cual se busca conocer o identificar sus preferencias está compuesta por los alumnos de los últimos años de educación secundaria (alumnos de los dos últimos años: 4to. y 5to. de secundaria). Donde se presupone que a este nivel los estudiantes ya tienen una idea, propia o inducida, sobre las preferencias hacia una determinada institución universitaria.

Bajo el supuesto de concebir a la elección de la universidad de preferencia como la elección de un “producto”; la aplicación práctica del presente estudio coincide plenamente con los objetivos y finalidades básicas propias del Análisis Conjunto. Esto es, determinar la contribución de cada atributo y niveles seleccionados en las preferencias globales de los encuestados (que para el caso son los estudiantes), y

establecer un modelo válido de los juicios de preferencia del consumidor que permitan realizar predicciones (Hair et al., 1999).

Algo importante, que complementa a lo expresado en el anterior párrafo, es que la evaluación de los atributos de la presente aplicación sigue un modelo compensatorio. Por lo tanto, justifica el uso del Análisis Conjunto como la técnica de análisis para la identificación de las preferencias universitarias.

### **3.3 SELECCIÓN DE ATRIBUTOS Y NIVELES**

Para la selección de atributos o factores consideraremos las recomendaciones teóricas para la selección de las mismas. Esto, considerando solo aquellos atributos determinantes que nos permitan explicar las preferencias de los encuestados.

En nuestra aplicación de identificación de preferencias universitarias, se tomó como principal referencia los atributos evaluados en el estudio de Soutar y Turner (2002), “Students' preferences for university: a conjoint analysis. *International Journal of Educational Management*”. Los factores del estudio referido así como otros considerados para la presente aplicación lo observamos en la tabla (III.1).

Así mismo, con fines de trabajar a nivel de la metodología AC tradicional se tuvo a bien seleccionar sólo 6 atributos del total de factores mencionados inicialmente. Al respecto, se debe mencionar que este proceso no siempre funciona así. Ya que, el número de atributos se determina en función de la necesidad del estudio. Y la metodología

conjunta a aplicar se elige posteriormente en función de la cantidad de atributos considerados.

Tabla III.1-Atributos o factores evaluados inicialmente

TIPO DE UNIVERSIDAD
DISTANCIA DESDE LA CASA
REPUTACIÓN ACADÉMICA
CALIDAD DE LA ENSEÑANZA
PERSPECTIVA DE EMPLEO
OPINIÓN DE LA FAMILIA
AMIGOS
ATMÓSFERA DEL CAMPUS DE ESTUDIO
APTITUD HACIA LOS CURSOS
INFRAESTRUCTURA ACADÉMICA
CONVENIOS/ALIANZAS
INFRAESTRUCTURA DEPORTIVA

FUENTE: Elaboración propia.

Para el proceso de selección de los atributos se utilizó la técnica cualitativa Grupos Focales. Los atributos seleccionados para la aplicación del presente estudio se pueden apreciar en la tabla (III.2).

Tabla III.2-Atributos o factores seleccionados

TIPO DE UNIVERSIDAD
REPUTACIÓN ACADÉMICA
CALIDAD DE LA ENSEÑANZA
INFRAESTRUCTURA ACADÉMICA
INFRAESTRUCTURA DEPORTIVA
PERSPECTIVA DE EMPLEO

FUENTE: Elaboración propia.

Así mismo, se obtuvieron los niveles para cada atributo seleccionado en la aplicación de identificación de preferencias universitarias. Esto lo observamos en la tabla (III.3).

Tabla III.3-Atributos y niveles seleccionados

ATRIBUTOS	NIVELES	
TIPO DE UNIVERSIDAD	Nacional	Particular
REPUTACIÓN ACADÉMICA	Si	No
CALIDAD DE LA ENSEÑANZA	Si	No
INFRAESTRUCTURA ACADÉMICA	Si	No
INFRAESTRUCTURA DEPORTIVA	Si	No
PERSPECTIVA DE EMPLEO	Si	No

FUENTE: Elaboración propia.

Los atributos considerados pueden ser descritos de la siguiente forma: **Tipo de universidad**, en referencia si la universidad es Pública (Nacional) o, si es Privada (Particular). **Reputación académica**, en alusión al prestigio sobresaliente con que cuenta una institución en todo ámbito. **Calidad de la enseñanza**, referido a la calidad en el servicio netamente educativo. **Infraestructura académica**, este atributo hace referencia a la implementación de laboratorios, salas de cómputo, bibliotecas, etc. Con que cuenta la institución universitaria. **Infraestructura deportiva**, a diferencia del anterior atributo, en este caso se valora el equipamiento y las instalaciones en el ámbito deportivo. Por último, en **Perspectivas de empleo**, se consideró a los programas de bolsas de empleo; así como, las perspectivas de trabajo de los egresados.

### 3.4 ESCOGIENDO LA METODOLOGÍA CONJUNTA

La metodología conjunta a utilizar depende principalmente del número de atributos incluidos en el estudio (Picón et al., 2006). En la presente investigación de identificación de preferencias universitarias se trabajó con 6 atributos. Por lo tanto, la metodología conjunta seleccionada fue el Análisis Conjunto Tradicional. A su vez,

elegir el Análisis Conjunto clásico implicó establecer un nivel de análisis individual y que la forma del modelo sea de tipo aditivo (Hair et al., 1999).

### 3.5 FORMAS DE PRESENTACIÓN

La forma de presentación escogida, a nivel del Análisis Conjunto tradicional, para el recojo de los datos siguió los siguientes criterios:

- **A nivel del tipo de combinaciones a presentar**, se estableció el procedimiento de Perfil Completo. Esto significa que se presentaron tarjetas conteniendo una combinación de todos los atributos con sus respectivos niveles.
- **A nivel de la forma de presentación de las combinaciones**, se establecieron descripciones verbales. Esto es, los perfiles se presentaron en tarjetas cuyo contenido mencionaban a los atributos y sus respectivos niveles.
- **A nivel de la forma de evaluar las preferencias sobre las combinaciones**, se establecieron puntuaciones que van desde el 1 hasta el 10. Donde 1 se consideró el menos preferido y 10, el más preferido.

### 3.6 CREACIÓN DE ESTÍMULOS

Como ya se ha mencionado la aplicación del presente estudio consideró 6 atributos. A su vez, cada atributo consideró sólo dos niveles (tabla III.3).

Por consiguiente, el número de todas las combinaciones en función a los atributos y niveles considerados se obtuvo del siguiente modo:  $N^{\circ}$  de Combinaciones =  $2^6 = 64$ . Esto es, se obtuvo 64 combinaciones (diseño factorial completo). El cual se traduce



considerando la metodología conjunta escogida, en 64 tarjetas que tendrían que ser evaluadas por cada encuestado. Esto implicaría una saturación o sobrecarga de información que puede suponer para los sujetos el estar evaluando muchos estímulos complejos a un mismo tiempo (Picón et al., 2006).

Así, los diversos perfiles se obtuvieron utilizando los comandos del programa estadístico R a través del package Conjoint (Bąk & Bartłomowicz, 2012).

Funciones R v. 3.1.1 - Diseño Factorial Completo

```
library(conjoint)
experiment = expand.grid(
  Tipo = c("Nacional", "Particular"),
  Repu = c("Si", "No"),
  Cali = c("Si", "No"),
  Acad = c("Si", "No"),
  Depo = c("Si", "No"),
  Empl = c("Si", "No"))
design<-caFactorialDesign(data=experiment,type="full")
print(design)
```

Obteniéndose, el diseño factorial completo con las 64 combinaciones de los atributos y niveles considerados en la investigación. El diseño factorial completo puede ser observado en la figura (III.1).

Debido al número elevado de combinaciones se procedió a considerar un Diseño Factorial Fraccionado. El cual permitió reducir esta primera cantidad de combinaciones hasta un menor número de tratamientos, pero, manteniendo la eficiencia a la hora de estimar los parámetros (Picón et al., 2006).

Figura III.1-Diseño Factorial Completo

	Tipo	Repu	Cali	Acad	Depo	Empl						
1	Nacional	Si	Si	Si	Si	Si	33	Nacional	Si	Si	Si	No
2	Particular	Si	Si	Si	Si	Si	34	Particular	Si	Si	Si	No
3	Nacional	No	Si	Si	Si	Si	35	Nacional	No	Si	Si	No
4	Particular	No	Si	Si	Si	Si	36	Particular	No	Si	Si	No
5	Nacional	Si	No	Si	Si	Si	37	Nacional	Si	No	Si	No
6	Particular	Si	No	Si	Si	Si	38	Particular	Si	No	Si	No
7	Nacional	No	No	Si	Si	Si	39	Nacional	No	No	Si	No
8	Particular	No	No	Si	Si	Si	40	Particular	No	No	Si	No
9	Nacional	Si	Si	No	Si	Si	41	Nacional	Si	Si	No	No
10	Particular	Si	Si	No	Si	Si	42	Particular	Si	Si	No	No
11	Nacional	No	Si	No	Si	Si	43	Nacional	No	Si	No	No
12	Particular	No	Si	No	Si	Si	44	Particular	No	Si	No	No
13	Nacional	Si	No	No	Si	Si	45	Nacional	Si	No	No	No
14	Particular	Si	No	No	Si	Si	46	Particular	Si	No	No	No
15	Nacional	No	No	No	Si	Si	47	Nacional	No	No	No	No
16	Particular	No	No	No	Si	Si	48	Particular	No	No	No	No
17	Nacional	Si	Si	Si	No	Si	49	Nacional	Si	Si	Si	No
18	Particular	Si	Si	Si	No	Si	50	Particular	Si	Si	Si	No
19	Nacional	No	Si	Si	No	Si	51	Nacional	No	Si	Si	No
20	Particular	No	Si	Si	No	Si	52	Particular	No	Si	Si	No
21	Nacional	Si	No	Si	No	Si	53	Nacional	Si	No	Si	No
22	Particular	Si	No	Si	No	Si	54	Particular	Si	No	Si	No
23	Nacional	No	No	Si	No	Si	55	Nacional	No	No	Si	No
24	Particular	No	No	Si	No	Si	56	Particular	No	No	Si	No
25	Nacional	Si	Si	No	No	Si	57	Nacional	Si	Si	No	No
26	Particular	Si	Si	No	No	Si	58	Particular	Si	Si	No	No
27	Nacional	No	Si	No	No	Si	59	Nacional	No	Si	No	No
28	Particular	No	Si	No	No	Si	60	Particular	No	Si	No	No
29	Nacional	Si	No	No	No	Si	61	Nacional	Si	No	No	No
30	Particular	Si	No	No	No	Si	62	Particular	Si	No	No	No
31	Nacional	No	No	No	No	Si	63	Nacional	No	No	No	No
32	Particular	No	No	No	No	Si	64	Particular	No	No	No	No

FUENTE: Resultados R v. 3.1.1

Para nuestro caso se utilizó el criterio de los Diseños Factoriales Fraccionadas Ortogonales. Pero, antes de determinar el diseño factorial fraccionado se encontró el número mínimo de combinaciones según lo establecido por Hair et al. (1999). Por lo tanto, teniendo un número total de niveles igual a 12 y un número de atributos igual a 6, el Número de combinaciones resultó igual a  $(12-6+1)$  obteniéndose un número mínimo de combinaciones igual a 7.

Luego, se procedió a construir el Diseño Factorial Fraccionado Ortogonal. Estos diseños se caracterizan por ser a la vez ortogonales y equilibrados. Para su elaboración se utilizó las funciones del programa estadístico R.

```
library(conjoint)
experiment = expand.grid(
  Tipo = c("Nacional", "Particular"),
  Repu = c("Si", "No"),
  Cali = c("Si", "No"),
  Acad = c("Si", "No"),
  Depo = c("Si", "No"),
  Empl = c("Si", "No"))
design<-caFactorialDesign(data=experiment, type="orthogonal")
print(design)

print(cor(caEncodedDesign(design)))
```

Figura III.2-Diseño Factorial Fraccionado Ortogonal

	Tipo	Repu	Cali	Acad	Depo	Empl
6	Particular	Si	No	Si	Si	Si
11	Nacional	No	Si	No	Si	Si
20	Particular	No	Si	Si	No	Si
29	Nacional	Si	No	No	No	Si
39	Nacional	No	No	Si	Si	No
42	Particular	Si	Si	No	Si	No
49	Nacional	Si	Si	Si	No	No
64	Particular	No	No	No	No	No

FUENTE: Resultados R v. 3.1.1

Los resultados obtenidos a través del aplicativo R nos proporcionaron 8 combinaciones (vale recordar que el número mínimo de estímulos obtenidos era de 7). En la figura (III.2) visualizamos la condición de “equilibrio” que deben cumplir los diseños factoriales fraccionados ortogonales. Donde cada nivel aparece dentro de cada atributo una cantidad igual de veces (Picón et al., 2006).

Así mismo, en la figura (III.3) se observa otra condición que implica la elección de un diseño factorial fraccionado ortogonal. Esto consiste en que las correlaciones de todos los atributos, considerados en el estudio, son iguales a cero.

Figura III.3-Correlaciones entre los atributos

	Tipo	Repu	Cali	Acad	Depo	Empl
Tipo	1	0	0	0	0	0
Repu	0	1	0	0	0	0
Cali	0	0	1	0	0	0
Acad	0	0	0	1	0	0
Depo	0	0	0	0	1	0
Empl	0	0	0	0	0	1

FUENTE: Resultados R v. 3.1.1

### 3.7 TRABAJO DE CAMPO Y TRATAMIENTO DE DATOS

Una vez obtenido el diseño factorial fraccionado ortogonal se procedió a crear las tarjetas siguiendo la forma de presentación preestablecida. Por lo tanto, se elaboró las ocho tarjetas donde cada una de ellas presentaba una combinación de los atributos con sus respectivos niveles en la forma de descripciones verbales y donde el encuestado tenía que elegir un puntaje de acuerdo a sus preferencias. El diseño de estas tarjetas se observan desde la figura (III.4) hasta la figura (III.11)

Figura III.4-Tarjeta N° 1 para el recojo de preferencias

Estudio sobre Preferencias Universitarias										
Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.										
TIPO DE UNIVERSIDAD:						Particular				
REPUTACIÓN ACADÉMICA:						Si				
CALIDAD DE LA ENSEÑANZA:						No				
INFRAESTRUCTURA ACADÉMICA:						Si				
INFRAESTRUCTURA DEPORTIVA:						Si				
PERSPECTIVAS DE EMPLEO:						Si				
1	2	3	4	5	6	7	8	9	10	
Muy poco preferida						Extremadamente preferida				

FUENTE: Elaboración propia

Figura III.5-Tarjeta N° 2 para el recojo de preferencias

Estudio sobre Preferencias Universitarias										
Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.										
TIPO DE UNIVERSIDAD:						Nacional				
REPUTACIÓN ACADÉMICA:						No				
CALIDAD DE LA ENSEÑANZA:						Si				
INFRAESTRUCTURA ACADÉMICA:						No				
INFRAESTRUCTURA DEPORTIVA:						Si				
PERSPECTIVAS DE EMPLEO:						Si				
1	2	3	4	5	6	7	8	9	10	
Muy poco preferida						Extremadamente preferida				

FUENTE: Elaboración propia

Figura III.6-Tarjeta N° 3 para el recojo de preferencias

Estudio sobre Preferencias Universitarias										
Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.										
TIPO DE UNIVERSIDAD:						Particular				
REPUTACIÓN ACADÉMICA:						No				
CALIDAD DE LA ENSEÑANZA:						Si				
INFRAESTRUCTURA ACADÉMICA:						Si				
INFRAESTRUCTURA DEPORTIVA:						No				
PERSPECTIVAS DE EMPLEO:						Si				
1	2	3	4	5	6	7	8	9	10	
Muy poco preferida						Extremadamente preferida				

FUENTE: Elaboración propia

Figura III.7-Tarjeta N° 4 para el recojo de preferencias

Estudio sobre Preferencias Universitarias										
Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.										
TIPO DE UNIVERSIDAD:						Nacional				
REPUTACIÓN ACADÉMICA:						Si				
CALIDAD DE LA ENSEÑANZA:						No				
INFRAESTRUCTURA ACADÉMICA:						No				
INFRAESTRUCTURA DEPORTIVA:						No				
PERSPECTIVAS DE EMPLEO:						Si				
1	2	3	4	5	6	7	8	9	10	
Muy poco preferida						Extremadamente preferida				

FUENTE: Elaboración propia

Figura III.8-Tarjeta N° 5 para el recojo de preferencias

Estudio sobre Preferencias Universitarias										
Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.										
TIPO DE UNIVERSIDAD:						Nacional				
REPUTACIÓN ACADÉMICA:						No				
CALIDAD DE LA ENSEÑANZA:						No				
INFRAESTRUCTURA ACADÉMICA:						Si				
INFRAESTRUCTURA DEPORTIVA:						Si				
PERSPECTIVAS DE EMPLEO:						No				
1	2	3	4	5	6	7	8	9	10	
Muy poco preferida						Extremadamente preferida				

FUENTE: Elaboración propia

Figura III.9-Tarjeta N° 6 para el recojo de preferencias

Estudio sobre Preferencias Universitarias										
Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.										
TIPO DE UNIVERSIDAD:						Particular				
REPUTACIÓN ACADÉMICA:						Si				
CALIDAD DE LA ENSEÑANZA:						Si				
INFRAESTRUCTURA ACADÉMICA:						No				
INFRAESTRUCTURA DEPORTIVA:						Si				
PERSPECTIVAS DE EMPLEO:						No				
1	2	3	4	5	6	7	8	9	10	
Muy poco preferida						Extremadamente preferida				

FUENTE: Elaboración propia

Figura III.10-Tarjeta N° 7 para el recojo de preferencias

Estudio sobre Preferencias Universitarias										
Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.										
TIPO DE UNIVERSIDAD:						Nacional				
REPUTACIÓN ACADÉMICA:						Si				
CALIDAD DE LA ENSEÑANZA:						Si				
INFRAESTRUCTURA ACADÉMICA:						Si				
INFRAESTRUCTURA DEPORTIVA:						No				
PERSPECTIVAS DE EMPLEO:						No				
1	2	3	4	5	6	7	8	9	10	
Muy poco preferida						Extremadamente preferida				

FUENTE: Elaboración propia



Figura III.11-Tarjeta N° 8 para el recojo de preferencias

<b>Estudio sobre Preferencias Universitarias</b>										
Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.										
<b>TIPO DE UNIVERSIDAD:</b>						Particular				
<b>REPUTACIÓN ACADÉMICA:</b>						No				
<b>CALIDAD DE LA ENSEÑANZA:</b>						No				
<b>INFRAESTRUCTURA ACADÉMICA:</b>						No				
<b>INFRAESTRUCTURA DEPORTIVA:</b>						No				
<b>PERSPECTIVAS DE EMPLEO:</b>						No				
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	
<b>Muy poco preferida</b>						<b>Extremadamente preferida</b>				

FUENTE: Elaboración propia

El trabajo de campo se planificó del siguiente modo: en primer lugar, el proceso de recolección de las preferencias se estableció haciendo entrega de una hoja que contenía las ochos tarjetas. Además, de establecer sus puntuaciones, cada alumno encuestado debía señalar su grado de instrucción, edad y su género. La hoja de recolección de preferencias podemos observarlo a nivel del Anexo 1. En segundo lugar, el tiempo establecido (media hora) para el llenado de la hoja se dio en un intervalo similar para ambos grados y su aplicación se realizó de forma simultánea.

Como ya se ha indicado, la presente investigación tiene por objetivo principal la aplicación de la técnica Análisis Conjunto tradicional en la identificación de preferencias. Por lo tanto, la muestra considerada, detallada en la parte metodológica, será suficiente para llevar a cabo la aplicación práctica de esta técnica.

El conjunto de estudiantes de 4to. y 5to. de nivel secundario considerados en el estudio presentaban las siguientes descripciones:

- El número de alumnos encuestados en total fueron 50 (40 en la primera entrevista y 10 adicionales para medir la fiabilidad temporal).

- De los cuales 27 alumnos pertenecían al 4to. Sec. y 23 al 5to. Sec.
- A nivel de las edades, el rango iba desde los 14 a los 18 años siendo los alumnos de 15 años los que presentaban una mayor frecuencia (22 estudiantes).
- Así mismo, a nivel del género, los registros indicaban 23 estudiantes mujeres y 27 estudiantes varones.

Una vez recogido los datos se procedieron al registro de los mismos a nivel del aplicativo R. Para este fin, y antes de la estimación de las utilidades a través de las funciones del package Conjoint, se generaron las siguientes tablas: en primer lugar, un vector donde se registraron de manera ordenada los nombres de los diversos niveles de cada atributo. A esta tabla se la denominó “*level*” (figura III.12). En segundo lugar, un vector donde se colocaron las diferentes puntuaciones de las preferencias (valores de la variable dependiente). A esta vector se la llamó “*prefe*” (figura III.13). Y por último, una matriz donde se registraron los códigos de los diversos niveles presentados en las combinaciones del diseño factorial fraccionado ortogonal. A esta matriz se la denominó con el nombre de “*code*” (figura III.14).

Figura III.12-Vector con los nombres de los niveles (“*level*”)

1	Nacional
2	Particular
3	Si
4	No
5	Si
6	No
7	Si
8	No
9	Si
10	No
11	Si
12	No

FUENTE: Resultados R v. 3.1.1



Figura III.13-Vector de preferencias (“*prefe*”)

1	7	36	2	246	5
2	6	37	4	247	5
3	8	38	7	248	1
4	5	39	6	249	5
5	4	40	3	250	8
6	6	41	7	251	7
7	8	42	8	252	3
8	2	43	9	253	2
9	4	44	8	254	6
10	6	45	6	255	7
11	9	46	9	256	1
12	2	47	7	257	6
13	2	48	5	258	7
14	8	49	6	259	6
15	7	50	8	260	2
16	1	51	9	261	4
17	5	52	4	262	5
18	8	53	2	263	6
19	6	54	7	264	1
20	4	55	7	265	2
21	3	56	1	266	7
22	2	57	3	267	8
23	7	58	2	268	2
24	1	59	8	269	1
25	7	60	1	270	6
26	5	61	1	271	6
27	9	62	5	272	1
28	5	63	4	273	8
29	5	64	1	274	5
30	6	65	1	275	7
31	6	66	1	276	5
32	1	67	9	277	6
33	7	68	1	278	8
34	4	69	1	279	5
35	9	70	7	280	7

FUENTE: Resultados R v. 3.1.1

En las segundas columnas de la figura (III.13) se observan las diversas puntuaciones, asignados por los alumnos, según las preferencias de los distintos perfiles evaluados.

Luego, a través del comando **caEncodedDesign** se obtuvo la matriz codificada del Diseño Factorial Fraccionado Ortogonal, observada en la figura (III.14). Tanto esta matriz como el vector con los nombres de los niveles y la matriz de preferencias nos permitieron obtener las estimaciones de las utilidades parciales y otros valores.

## Funciones R v. 3.1.1 – Codificación de los niveles

```

library(conjoint)
experiment = expand.grid(
  Tipo = c("Nacional", "Particular"),
  Repu = c("Si", "No"),
  Cali = c("Si", "No"),
  Acad = c("Si", "No"),
  Depo = c("Si", "No"),
  Empl = c("Si", "No"))
design<-caFactorialDesign(data=experiment, type="orthogonal")
print(design)

code<-caEncodedDesign(design)
print(code)

```

Figura III.14-Matriz de perfiles codificados (“code”)

	Tipo	Repu	Cali	Acad	Depo	Empl
6	2	1	2	1	1	1
11	1	2	1	2	1	1
20	2	2	1	1	2	1
29	1	1	2	2	2	1
39	1	2	2	1	1	2
42	2	1	1	2	1	2
49	1	1	1	1	2	2
64	2	2	2	2	2	2

FUENTE: Resultados R v. 3.1.1

Una observación importante a señalar es que tanto el vector de preferencias visto en la figura (III.13) como la matriz de perfiles codificados visto en la figura (III.14) no requieren ser convertidos a matrices para la estimación de utilidades como para cualquier otro análisis a través de las funciones del package Conjoint. Al momento de su lectura a través del programa R bastará, por ejemplo, con el uso de la función **read.table** con el nombre de los archivos desde su respectiva ubicación.

### 3.8 ESTIMACIÓN DE LAS UTILIDADES

Una vez elaborado las diversas tablas para su ejecución en el programa estadístico R, se procedió a realizar los diversos análisis estadísticos utilizando las funciones del package Conjoint (Bak & Bartłomowicz, 2012).

En el año 2009 Andrzej Bak y Tomasz Bartłomowicz investigadores del Departamento de Econometría y Ciencias de la Computación en la Wroclaw University of Economics dan a conocer, a través de su publicación “Conjoint analysis method and its implementation in conjoint R package”, el desarrollo del paquete Conjoint<sup>3</sup> para el programa estadístico R.

Tabla III.4-Funciones del conjoint R package

FUNCIONES	DESCRIPCIÓN
caPartUtilities(y, x, z)	la función calcula matriz de utilidades de los niveles individuales de los encuestados (con intercepto en primer lugar)
caTotalUtilities(y, x)	la función calcula la matriz de utilidades total teóricos para los encuestados (para los n perfiles y todos los encuestados)
caImportance(y, x)	calcula la importancia de todos los atributos. La suma de la importancia debe ser 100%
caUtilities(y, x, z)	la función calcula las utilidades de los niveles del atributo
Conjoint(y, x, z)	la función devuelve el valor de las utilidades parciales de los niveles (los parámetros del modelo son para toda la muestra), el vector de porcentaje de la importancia de los atributos, la suma de ellos y sus correspondientes gráficos (gráficos de barras). La suma de la importancia debe ser 100%
ShowAllUtilities(y, x, z)	la función devuelve la matriz básica de las utilidades parciales (con el intercepto), la matriz de utilidades totales (para los n perfiles y todos los encuestados) y el vector de porcentajes de la importancia de los atributos, con la suma de ellos. La suma de la importancia debe ser 100%
caBTL(sym, y, x)	la función calcula la participación de los perfiles de simulación mediante el modelo probabilístico BTL (Bradley-Terry-Luce). La suma de la participación debe ser 100%
caLogit(sym, y, x)	la función calcula la participación de los perfiles de simulación mediante el modelo logit. La suma de la participación debe ser 100%
caMaxUtility(sym, y, x)	la función calcula la participación de los perfiles de simulación mediante modelo de máxima utilidad ("primera posición"). La suma de la participación debe ser 100%
ShowAllSimulations(sym, y, x)	la función devuelve 3 vectores de participaciones en porcentajes utilizando la máxima utilidad, de los modelos BTL y logit. La suma de importancia para cada vector debe ser 100%
caModel(y, x)	la función calcula los parámetros del modelo análisis conjunto
caSegmentation(y, x, c=3)	la función divide a los encuestados en tres o n clusters usando el método k-means - función kmeans() del paquete stats
Function's arguments	
y	matriz (vector) de preferencias empíricas
x	matriz de perfiles
z	vector con los nombres de los niveles
sym	matriz de perfiles de simulación
c	número de clusters (opcional), valor por defecto: c = 3

FUENTE: Bak and Bartłomowicz (2009)

<sup>3</sup> <https://cran.r-project.org/web/packages/conjoint/index.html>

Lo importante de este aporte es que el Conjoint R Package es fácil de utilizar y está basado en el Análisis Conjunto tradicional con el método de recogida de Perfil Completo y el modelo de Regresión Lineal Múltiple con variables ficticias (dummy) (Bak & Bartlomowicz, 2009).

En la tabla (III.4) se describen las funciones principales que presenta el package Conjoint incluyendo los argumentos necesarios para su ejecución. Estas complementan las funciones ya utilizadas, en el estudio, del mismo paquete: **caFactorialDesign** y **caEncodedDesign**. Para la presente aplicación no será necesario el uso de todas las funciones del Conjoint R Package descritas.

Como primer paso se procedió con el análisis que determinó la Matriz de utilidades de los niveles individuales de los encuestados (con la estimación del intercepto en primer lugar) haciendo uso de la función **caPartUtilities**.

Funciones R v. 3.1.1 – Matriz de utilidades de los niveles

<code>caPartUtilities(y=prefe, x=code, z=level)</code>
--

En el argumento de la función **caPartUtilities** se consideró al vector de preferencias (*prefe*), la matriz de perfiles codificados (*code*) y el vector que incluye los nombres de los niveles de cada atributo (*level*). Vistos en las figuras (III.13), (III.14) y (III.12) respectivamente.

Figura III.15-Matriz de utilidades de los niveles

	intercept	Nacional	Particular	Si	No	Si	No	Si	No	Si	No	Si	No
[1,]	5.750	0.000	0.000	0.750	-0.750	1.250	-1.250	1.000	-1.000	0.000	0.000	0.750	-0.750
[2,]	4.875	-0.625	0.625	0.375	-0.375	2.625	-2.625	0.625	-0.625	0.125	-0.125	0.375	-0.375
[3,]	4.500	1.000	-1.000	0.000	0.000	1.250	-1.250	0.750	-0.750	0.000	0.000	1.250	-1.250
[4,]	5.500	-0.250	0.250	0.500	-0.500	1.000	-1.000	1.250	-1.250	0.250	-0.250	1.000	-1.000
[5,]	5.250	-1.250	1.250	0.250	-0.250	1.250	-1.250	1.250	-1.250	0.250	-0.250	0.250	-0.250
[6,]	7.375	-0.125	0.125	0.375	-0.375	0.875	-0.875	-0.125	0.125	0.125	-0.125	0.625	-0.625
[7,]	5.500	-0.250	0.250	0.500	-0.500	2.250	-2.250	0.500	-0.500	0.250	-0.250	1.250	-1.250
[8,]	3.125	-1.125	1.125	0.125	-0.125	1.625	-1.625	0.875	-0.875	-0.375	0.375	0.375	-0.375
[9,]	2.875	-1.625	1.625	-0.125	0.125	1.875	-1.875	0.375	-0.375	-0.375	0.375	0.125	-0.125
[10,]	3.625	0.375	-0.375	1.375	-1.375	1.625	-1.625	0.125	-0.125	0.375	-0.375	-0.875	0.875
[11,]	6.875	0.875	-0.875	0.875	-0.875	1.125	-1.125	-0.375	0.375	0.125	-0.125	-0.625	0.625
[12,]	3.750	-0.500	0.500	0.750	-0.750	1.250	-1.250	1.000	-1.000	0.000	0.000	0.250	-0.250
[13,]	4.125	-0.375	0.375	1.125	-1.125	0.125	-0.125	0.125	-0.125	0.875	-0.875	1.375	-1.375
[14,]	5.500	1.000	-1.000	1.500	-1.500	1.250	-1.250	0.250	-0.250	0.000	0.000	0.250	-0.250
[15,]	4.375	-0.125	0.125	0.625	-0.625	0.875	-0.875	1.125	-1.125	0.375	-0.375	0.875	-0.875
[16,]	6.375	-0.125	0.125	0.875	-0.875	1.625	-1.625	1.125	-1.125	0.625	-0.625	1.375	-1.375
[17,]	4.750	0.250	-0.250	0.000	0.000	1.500	-1.500	0.750	-0.750	0.000	0.000	0.750	-0.750
[18,]	3.375	0.125	-0.125	0.125	-0.125	0.875	-0.875	0.375	-0.375	1.125	-1.125	0.375	-0.375
[19,]	5.125	1.625	-1.625	0.875	-0.875	0.375	-0.375	0.625	-0.625	0.125	-0.125	0.875	-0.875
[20,]	4.750	0.250	-0.250	0.750	-0.750	1.000	-1.000	0.500	-0.500	0.750	-0.750	0.500	-0.500
[21,]	3.750	0.250	-0.250	0.750	-0.750	2.250	-2.250	0.250	-0.250	-0.250	0.250	-0.250	0.250
[22,]	2.750	0.000	0.000	1.000	-1.000	1.250	-1.250	0.250	-0.250	-0.250	0.250	-0.500	0.500
[23,]	3.875	-0.625	0.625	0.375	-0.375	1.125	-1.125	1.125	-1.125	0.125	-0.125	0.875	-0.875
[24,]	5.625	-0.625	0.625	0.875	-0.875	1.125	-1.125	0.625	-0.625	1.875	-1.875	0.125	-0.125
[25,]	3.625	0.125	-0.125	-1.375	1.375	1.875	-1.875	0.875	-0.875	-0.125	0.125	1.625	-1.625
[26,]	4.375	0.125	-0.125	-0.125	0.125	1.125	-1.125	-0.125	0.125	-0.125	0.125	1.875	-1.875
[27,]	5.500	0.000	0.000	0.750	-0.750	1.750	-1.750	0.500	-0.500	-0.250	0.250	0.000	0.000
[28,]	3.750	0.000	0.000	0.500	-0.500	0.750	-0.750	0.750	-0.750	-0.250	0.250	0.500	-0.500
[29,]	2.000	-0.500	0.500	0.250	-0.250	0.000	0.000	0.250	-0.250	0.250	-0.250	0.250	-0.250
[30,]	3.125	0.625	-0.625	-1.125	1.125	1.125	-1.125	-1.125	1.125	1.125	-1.125	0.125	-0.125
[31,]	3.750	0.000	0.000	0.750	-0.750	0.750	-0.750	0.500	-0.500	0.000	0.000	0.250	-0.250
[32,]	4.875	0.125	-0.125	0.375	-0.375	2.125	-2.125	0.375	-0.375	0.375	-0.375	0.875	-0.875
[33,]	4.625	0.125	-0.125	0.125	-0.125	1.375	-1.375	0.875	-0.875	0.875	-0.875	0.625	-0.625
[34,]	4.125	-0.125	0.125	-0.125	0.125	2.625	-2.625	0.125	-0.125	-0.125	0.125	0.625	-0.625
[35,]	6.375	-1.125	1.125	0.125	-0.125	-0.125	0.125	0.125	-0.125	0.375	-0.375	-0.125	0.125

FUENTE: Resultados R v. 3.1.1

En la figura (III.15) se observan las estimaciones de las utilidades de cada nivel (utilidades parciales) por cada uno de los encuestados, incluyendo sus interceptos.

De forma seguida se procedió a calcular las importancias relativas de cada uno de los atributos considerados en el estudio. Siempre, considerando al total de encuestados (estudiantes).

Funciones R v. 3.1.1 – Importancia de los atributos

```
imp<-caImportance(y=prefe, x=code)
print("Importance summary: ", quote=FALSE)
print(imp)
print(paste("Sum: ", sum(imp)), quote=FALSE)
```

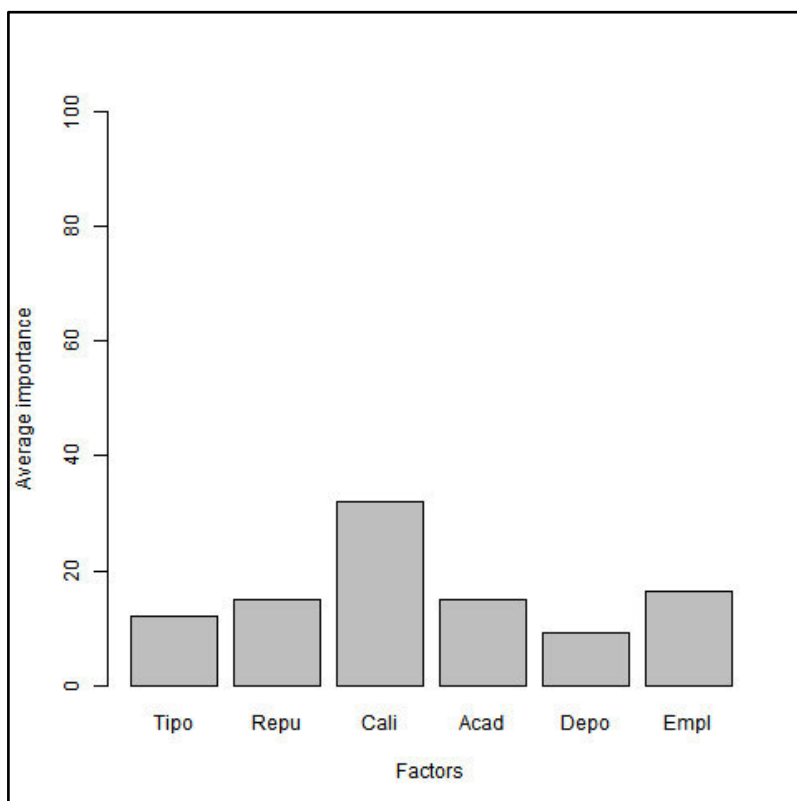
Figura III.16-Importancia relativa de los atributos

```
[1] Importance summary:
> print(imp)
[1] 12.21 14.91 32.06 15.14  9.16 16.52
> print(paste("Sum: ", sum(imp)), quote=FALSE)
[1] Sum: 100
```

FUENTE: Resultados R v. 3.1.1

Los resultados obtenidos, considerando al total de estudiantes, otorgan al atributo Calidad de la Enseñanza (32,06%) la mayor importancia, seguido de la Perspectiva de Empleo (16,52%), Infraestructura Académica (15,14%), Reputación Académica (14,91%), Tipo de Universidad (12,21%) y la Infraestructura Deportiva (9,16%).

Figura III.17-Importancia relativa de los atributos



FUENTE: Resultados R v. 3.1.1

Luego, se procedió a calcular las utilidades de los niveles de los atributos, incluido el intercepto, utilizando la función **caUtilities**. A diferencia de los resultados obtenidos con la función **caPartUtilities**, la función **caUtilities** obtiene las estimaciones considerando a todos los individuos. Lo que nos permite conocer las preferencias por nivel de cada uno de los atributos considerados.

Funciones R v. 3.1.1 – Utilidades Parciales

```
caUtilities(y=prefe, x=code, z=level)
```

Figura III.18-Estimación de las utilidades parciales

```
Call:
lm(formula = frml)

Residuals:
    Min       1Q   Median       3Q      Max
-4,5929 -1,0214 -0,0214  1,4500  5,3214

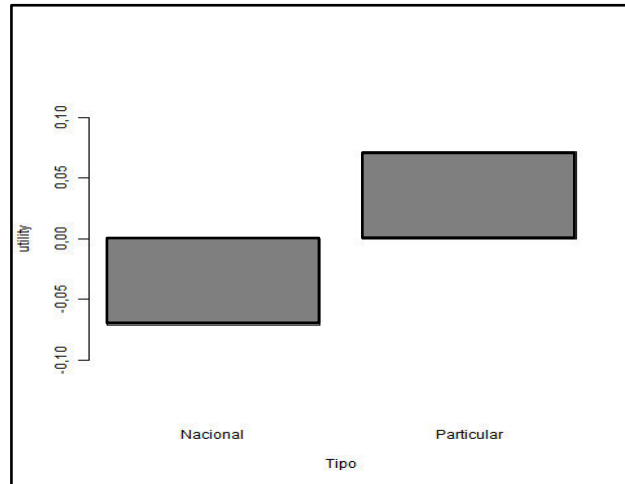
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    4,55714    0,11668   39,058 < 2e-16 ***
factor(x$Tipo)1 -0,07143    0,11668   -0,612  0,540915
factor(x$Repu)1  0,42143    0,11668    3,612  0,000362 ***
factor(x$Cali)1  1,27857    0,11668   10,958 < 2e-16 ***
factor(x$Acad)1  0,50000    0,11668    4,285  2,53e-05 ***
factor(x$Depo)1  0,23571    0,11668    2,020  0,044333 *
factor(x$Empl)1  0,51429    0,11668    4,408  1,50e-05 ***
---
Signif. codes:  0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Residual standard error: 1,952 on 273 degrees of freedom
Multiple R-squared:  0.3911,    Adjusted R-squared:  0.3778
F-statistic: 29.23 on 6 and 273 DF,  p-value: < 2,2e-16

[1] 4.55714286 -0.07142857 0.07142857 0.42142857 -0.42142857 1.27857143
[7] -1.27857143 0.50000000 -0.50000000 0.23571429 -0.23571429 0.51428571
[13] -0.51428571
```

FUENTE: Resultados R v. 3.1.1

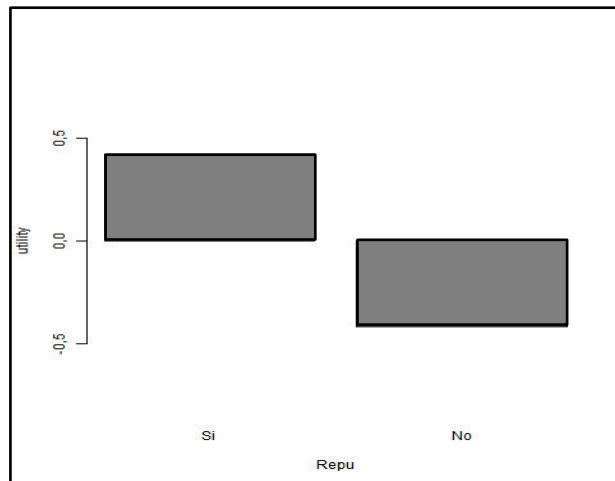
Figura III.19-Tipo de universidad: preferencia de los niveles



FUENTE: Resultados R v. 3.1.1

De los niveles considerados en el atributo **Tipo de Universidad** (*Nacional* o *Particular*), son las universidades particulares en comparación a las universidades nacionales, los que presentaron una mayor preferencia por parte del total de estudiantes encuestados.

Figura III.20-Reputación académica: preferencia de los niveles

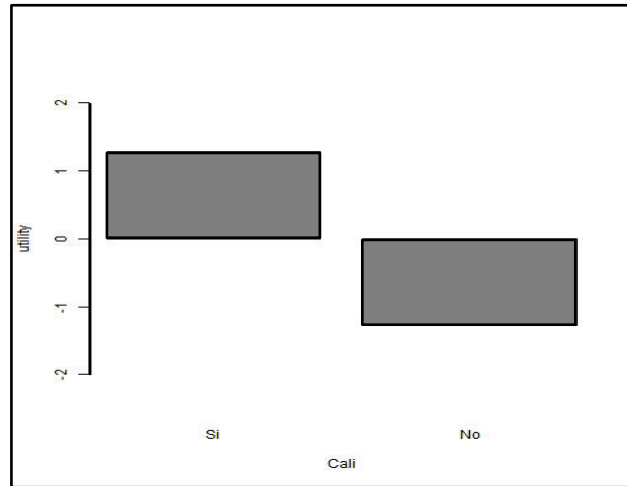


FUENTE: Resultados R v. 3.1.1

De los niveles considerados en el atributo **Reputación Académica** (*Si* o *No*), los alumnos encuestados, valoraron más la presencia de la **Reputación Académica** al momento de calificar sus preferencias universitarias.



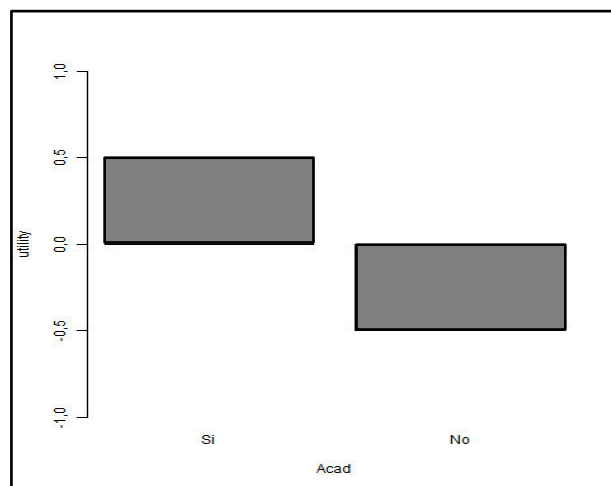
Figura III.21-Calidad de la enseñanza: preferencia de los niveles



FUENTE: Resultados R v. 3.1.1

De los niveles considerados en el atributo **Calidad de la Enseñanza** (*Si* o *No*), la presencia de una adecuada Calidad de la Enseñanza fue valorado más por los alumnos encuestados en la calificación de sus preferencia universitarias.

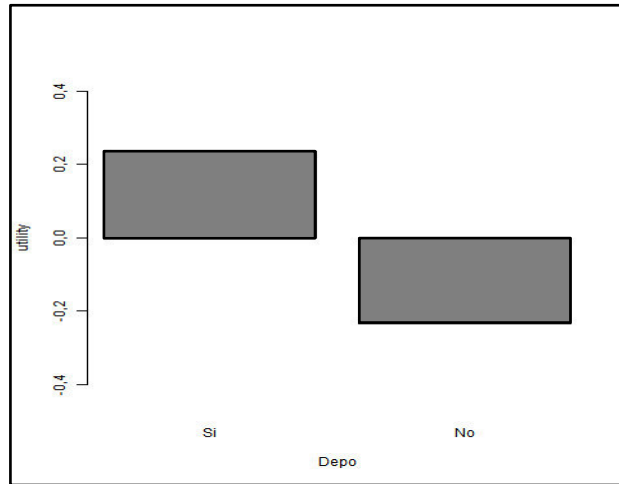
Figura III.22-Infraestructura académica: preferencia de los niveles



FUENTE: Resultados R v. 3.1.1

De los niveles considerados en el atributo **Infraestructura Académica** (*Si* o *No*), los encuestados valoraron más que la institución universitaria de su preferencia cuente con una adecuada **Infraestructura Académica**.

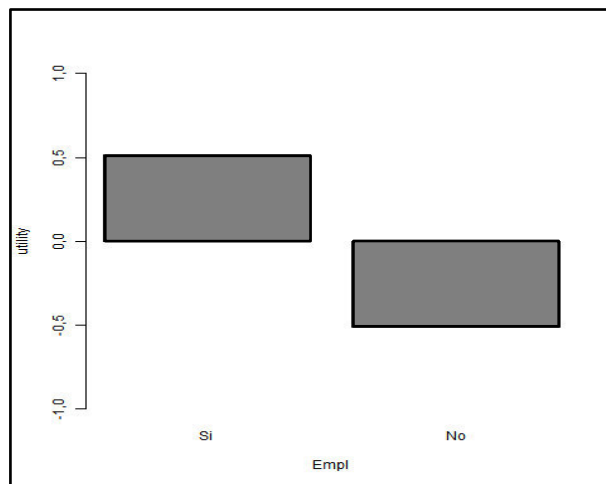
Figura III.23-Infraestructura deportiva: preferencia de los niveles



FUENTE: Resultados R v. 3.1.1

De los niveles considerados en el atributo **Infraestructura Deportiva** (*Si* o *No*), los encuestados valoraron más que la institución universitaria de su preferencia cuente con una adecuada **Infraestructura Deportiva**.

Figura III.24-Perspectiva de empleo: preferencia de los niveles



FUENTE: Resultados R v. 3.1.1

De los niveles considerados en el atributo **Perspectiva de Empleo** (*Si* o *No*), los encuestados valoraron más el hecho de que la universidad de su preferencia les garantice un empleo culminado sus estudios.

La función Conjoint devolvió el valor de las utilidades parciales (**part-worths**) de los niveles (los parámetros del modelo son para toda la muestra), el vector de porcentaje de la importancia de los atributos, la suma de ellos y sus correspondientes gráficos (gráficos de barras). La suma de la importancia debe ser 100%.

Funciones R v. 3.1.1 – Estimación General

Conjoint(y=prefe, x=code, z=level)

Figura III.25-Estimación General

```
Call:
lm(formula = frml)

Residuals:
    Min       1Q   Median       3Q      Max
-4,5929 -1,0214 -0,0214  1,4500  5,3214

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    4,55714    0,11668   39,058  < 2e-16 ***
factor(x$Tipo)1 -0,07143    0,11668   -0,612  0,540915
factor(x$Repu)1  0,42143    0,11668    3,612  0,000362 ***
factor(x$Cali)1  1,27857    0,11668   10,958  < 2e-16 ***
factor(x$Acad)1  0,50000    0,11668    4,285  2,53e-05 ***
factor(x$Depo)1  0,23571    0,11668    2,020  0,044333 *
factor(x$Empl)1  0,51429    0,11668    4,408  1,50e-05 ***
---
Signif. codes:  0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Residual standard error: 1,952 on 273 degrees of freedom
Multiple R-squared:  0.3911,    Adjusted R-squared:  0.3778
F-statistic: 29.23 on 6 and 273 DF,  p-value: < 2,2e-16

[1] "Part worths (utilities) of levels (model parameters for whole sample):"
      levnms      utls
1  intercept  4,5571
2  Nacional  -0,0714
3  Particular 0,0714
4      Si    0,4214
5      No   -0,4214
6      Si    1,2786
7      No   -1,2786
8      Si     0,5
9      No    -0,5
10     Si    0,2357
11     No   -0,2357
12     Si    0,5143
13     No   -0,5143
[1] "Average importance of factors (attributes):"
[1] 12,21 14,91 32,06 15,14  9,16 16,52
[1] Sum of average importance: 100
[1] "Chart of average factors importance"
```

FUENTE: Resultados R v. 3.1.1

Los resultados obtenidos como consecuencia de la ejecución de la función Conjoint nos indicaron en primer lugar, un análisis descriptivo de los residuales (máximo y mínimo valor, los valores del primer y tercer cuartil, y la mediana). Así mismo, se presentaba las estimaciones de los parámetros del modelo AC y sus respectivas significancias. En función a estos últimos resultados podemos construir el siguiente modelo de regresión con variables ficticias (dummy):

$$\begin{aligned}\hat{U}_i = & 4,5571 - 0,0714x_{11} + 0,0714x_{12} + 0,4214x_{21} - 0,4214x_{21} + 1,2786x_{31} - 1,2786x_{32} \\ & + 0,5x_{41} - 0,5x_{42} + 0,2357x_{51} - 0,2357x_{51} + 0,5143x_{61} - 0,5143x_{61}\end{aligned}$$

*Donde:*

$x_{jk}$ : Puede asumir valores 0: ausencia o 1: presencia en el  $k$ -ésimo nivel del  $j$ -ésimo atributo

A nivel de la bondad de ajuste del modelo se observó el valor del estadístico F con un p\_valor menor a 0,05. Rechazando la hipótesis nula de que todos los parámetros del modelo son iguales a 0. También, las estimaciones obtenidas a nivel individual de los parámetros en función al estadístico t, nos mostró valores significativos al 5% en casi todos los parámetros estimados incluido el intercepto. La única excepción lo constituyó el factor Tipo de Universidad (p\_valor=0,54)

Como se observa a través de la ejecución de la función **Conjoint** también se obtienen los resultados de las utilidades parciales de los niveles (resultados obtenidos cuando se utilizó la función **caUtilities**), y las importancias relativas de los atributos (valores obtenidos cuando se utilizó la función **caImportance**).

### 3.9 EVALUACIÓN DEL AJUSTE

La evaluación del ajuste de un modelo de Análisis Conjunto se da principalmente a través de los siguientes tres criterios:

- Evaluación de la fiabilidad.
- Validación interna.
- Validación predictiva.

La evaluación de la **fiabilidad** se obtuvo en función a la correlación producto-momento de Pearson entre las utilidades (normalizadas) obtenidas de la evaluación principal y las obtenidas de la evaluación para testar la fiabilidad (Picón et al., 2006).

En este sentido, se procedió en primer lugar, a normalizar las utilidades parciales estimadas obtenidas de la evaluación principal. Esto es, considerando la matriz de preferencias universitarias (ver Anexo 2).

Los resultados obtenidos se observan en la figura (III.26).

Figura III.26-Utilidades normalizadas de la evaluación principal

	V1
[1,]	2.972353801
[2,]	-0.349340893
[3,]	-0.246858879
[4,]	0.004322528
[5,]	0.004322528
[6,]	0.619501677
[7,]	-1.215701449
[8,]	0.060730696
[9,]	-0.656930467
[10,]	-0.128947150
[11,]	-0.467252622
[12,]	0.070993250
[13,]	-0.667193022

FUENTE: Resultados R v. 3.1.1

Posteriormente se procedió a estimar las utilidades parciales normalizadas de las puntuaciones de los perfiles reservados para evaluar la fiabilidad temporal (Anexo 4).

Figura III.27-Utilidades normalizadas de la evaluación complementaria

	V1
[1,]	2.93013772
[2,]	-0.35204325
[3,]	-0.13631304
[4,]	0.14105436
[5,]	-0.62941065
[6,]	0.49546827
[7,]	-0.98382456
[8,]	0.40301247
[9,]	-0.89136876
[10,]	-0.09008514
[11,]	-0.39827115
[12,]	0.20269157
[13,]	-0.69104785

FUENTE: Resultados R v. 3.1.1

La correlación entre las utilidades normalizadas de la evaluación principal y las utilidades normalizadas de la evaluación complementaria se realizó a través del coeficiente de Pearson obteniéndose un valor de  $r=0,97$ .

El valor obtenido muy cercano a 1 ( $r=0,97$ ) indica una fuerte asociación. Y se puede interpretar que existe un alto grado de precisión de las utilidades estimadas.

Las medidas de fiabilidad se complementan con las medidas de validez de los resultados (Green & Rao, 1971) y buscan responder las siguiente interrogante ¿La estimación de los parámetros para la ecuación representativa del análisis conjunto permite derivar preferencias que reproducen fielmente las ofrecidas por la muestra de entrevistados? (Casielles, 1990).

La validación interna indica el grado de ajuste entre las preferencias (utilidades) estimadas por el análisis conjunto y las que muestra de hecho el sujeto (Picón et al., 2006). Su cuantificación se realiza a través del coeficiente de determinación  $R^2$  de Pearson o Tau de Kendall. La elección del coeficiente depende de la escala en la que se midió la variable dependiente (ordinal o de intervalo). En el presente estudio la variable dependiente recogió puntuaciones del 1 al 10; por lo tanto, la validación interna se evaluó en función del coeficiente de determinación  $R^2$ .

Si se considera el coeficiente de determinación obtenido en la Figura III.25 ( $R^2=0,39$ ) podríamos concluir que existe un ajuste bajo. Pero, con respecto a este punto, habría que señalar que la aplicación del método AC tradicional está orientada a un nivel de análisis de tipo individual y no agregado. Por lo tanto, al obtener las estimaciones agregadas, sin efectuar previamente una segmentación, considerando al conjunto de encuestados total no sorprende obtener resultados muy pobres a nivel del  $R^2$  y a nivel de la capacidad predictiva (Hair et al., 1999, p. 435).

Lo expresado en el anterior párrafo se verifica al realizar el análisis por cada estudiante encuestado. Así, si consideramos al estudiante 2 y hacemos el análisis de sus preferencias obtenemos un coeficiente de determinación  $R^2=0,9827$  (ver Figura III.28).

Para realizar la estimación individual se utilizó el comando `caModel` evaluando la matriz de preferencias universitarias registrado en el Anexo 2.

Funciones R v. 3.1.1 – Estimación individual

---

```
caModel(y=prefem[2,], x=code)
```

---

Figura III.28-Estimación individual – encuestado n° 2

```

Call:
lm(formula = frml)

Residuals:
    1     2     3     4     5     6     7     8 
-0.375 -0.375  0.375  0.375  0.375  0.375 -0.375 -0.375 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      4.875      0.375   13.000   0.0489 *
factor(x$Tipo)1  -0.625      0.375   -1.667   0.3440
factor(x$Repu)1    0.375      0.375    1.000   0.5000
factor(x$Cali)1    2.625      0.375    7.000   0.0903 .
factor(x$Acad)1    0.625      0.375    1.667   0.3440
factor(x$Depo)1    0.125      0.375    0.333   0.7952
factor(x$Empl)1    0.375      0.375    1.000   0.5000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.061 on 1 degrees of freedom
Multiple R-squared: 0.9827,    Adjusted R-squared: 0.8786 
F-statistic: 9.444 on 6 and 1 DF,  p-value: 0.2441

```

FUENTE: Resultados R v. 3.1.1

La validación cruzada se utiliza como criterio para evaluar el poder predictivo del modelo. Al respecto, Rodolfo Casielles (1990) señala que “los tests de validez externa trabajan con dos submuestras cada una indicando preferencias por perfiles de productos. Una submuestra se utiliza para estimar los partworths. Posteriormente se investiga si con dichos partworths se pueden pronosticar las preferencias de la segunda submuestra” (p. 19). Esta segunda submuestra se conforma de un conjunto de perfiles de reserva (estímulos holdout).

En el estudio se reservó una submuestra correspondiente a 5 encuestados (ver Anexo 3). Estos perfiles no fueron considerados en la evaluación principal para la estimación de las utilidades del modelo.



El modelo considerado para la estimación de las utilidades totales (preferencias) en la evaluación principal fue el siguiente:

$$\begin{aligned}\hat{U}_i = & 4,5571 - 0,0714x_{11} + 0,0714x_{12} + 0,4214x_{21} - 0,4214x_{21} + 1,2786x_{31} - 1,2786x_{32} \\ & + 0,5x_{41} - 0,5x_{42} + 0,2357x_{51} - 0,2357x_{51} + 0,5143x_{61} - 0,5143x_{61}\end{aligned}$$

Donde:

$$x_{jk} = 0 \text{ ó } 1 \quad (j: 1,2 \dots, 6)$$

Las estimaciones de las utilidades totales fueron dadas en función a la evaluación de los modelos observados en la tabla (III.5).

Tabla III.5-Estimación utilidades totales del modelo AC

5,0214	=	4,5571 + 0,0714(1) + 0,4214(1) - 1,2786(1) + 0,5(1) + 0,2357(1) + 0,5143(1)
5,5929	=	4,5571 - 0,0714(1) - 0,4214(1) + 1,2786(1) - 0,5(1) + 0,2357(1) + 0,5143(1)
6,2643	=	4,5571 + 0,0714(1) - 0,4214(1) + 1,2786(1) + 0,5(1) - 0,2357(1) + 0,5143(1)
3,4072	=	4,5571 - 0,0714(1) + 0,4214(1) - 1,2786(1) - 0,5(1) - 0,2357(1) + 0,5143(1)
3,0071	=	4,5571 - 0,0714(1) - 0,4214(1) - 1,2786(1) + 0,5(1) + 0,2357(1) - 0,5143(1)
5,5500	=	4,5571 + 0,0714(1) + 0,4214(1) + 1,2786(1) - 0,5(1) + 0,2357(1) - 0,5143(1)
5,9357	=	4,5571 - 0,0714(1) + 0,4214(1) + 1,2786(1) + 0,5(1) - 0,2357(1) - 0,5143(1)
1,6786	=	4,5571 + 0,0714(1) - 0,4214(1) - 1,2786(1) - 0,5(1) - 0,2357(1) - 0,5143(1)

FUENTE: Elaboración propia

Una vez obtenidos las estimaciones utilidades totales (UT) en la evaluación principal se procedió a relacionarlas con las estimaciones de las utilidades totales obtenidos en las encuestas de reserva o holdout (ver anexo 3) a través de la correlación de Pearson (Rao, 2014, p. 64).

## Funciones R v. 3.1.1 – Estimación utilidades totales de los perfiles de reserva

```
holdout<-caTotalUtilities(y=holdout, x=code)
colMeans(caTotalUtilities(y=holdout, x=code))
```

Figura III.29-Estimación utilidades totales de los perfiles de reserva

```
> holdout
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
[1,] 3.625 7.625 8.375 2.375 1.375 9.375 9.625 0.625
[2,] 4.750 6.750 6.250 2.250 1.250 9.250 6.750 0.750
[3,] 3.750 5.750 7.250 2.250 3.250 3.250 5.750 0.750
[4,] 4.875 4.875 7.125 3.125 2.125 2.125 3.875 0.875
[5,] 2.500 3.500 0.500 0.500 3.500 5.500 2.500 1.500
>
> colMeans(caTotalUtilities(y=holdout, x=code))
[1] 3.9 5.7 5.9 2.1 2.3 5.9 5.7 0.9
```

FUENTE: Resultados R v. 3.1.1

La correlación obtenida  $r=0,97$  nos indicó una muy buena capacidad predictiva del modelo conjunto propuesto.

### 3.10 INTERPRETACIÓN DE LOS RESULTADOS

A nivel de un análisis agregado, la combinación ideal estuvo conformada por aquellos niveles que presentaron una mayor utilidad parcial. Esta combinación se presentó en la tabla (III.6).

Tabla III.6-Combinación ideal

ATRIBUTOS	NIVELES	
TIPO DE UNIVERSIDAD	Nacional	Particular
REPUTACIÓN ACADÉMICA	Si	No
CALIDAD DE LA ENSEÑANZA	Si	No
INFRAESTRUCTURA ACADÉMICA	Si	No
INFRAESTRUCTURA DEPORTIVA	Si	No
PERSPECTIVA DE EMPLEO	Si	No

FUENTE: Elaboración propia

Luego, los atributos más importantes según el nivel de importancia relativa, presentado en la figura (III.17), seguían el siguiente orden:

- 1°. Calidad de la Enseñanza (32,06%)
- 2°. Perspectiva de Empleo (16,52%)
- 3°. Infraestructura Académica (15,14%)
- 4°. Reputación Académica (14,91%)
- 5°. Tipo de Universidad (12,21%)
- 6°. La Infraestructura Deportiva (9,16%)

También, se aprecia en la tabla (III.7) que el perfil ortogonal más preferido por los encuestados fue la presentada en la tarjeta N° 3 (figura III.30). Esta presenta una utilidad total estimada de 6,24.

Figura III.30-Estímulo con mayor preferencia

**Estudio sobre Preferencias Universitarias**

Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.

TIPO DE UNIVERSIDAD:	Particular
REPUTACIÓN ACADÉMICA:	No
CALIDAD DE LA ENSEÑANZA:	Si
INFRAESTRUCTURA ACADÉMICA:	Si
INFRAESTRUCTURA DEPORTIVA:	No
PERSPECTIVAS DE EMPLEO:	Si

12345678910

Muy poco preferida
Extremadamente preferida

FUENTE: Elaboración propia

Tabla III.7-Utilidades totales de las tarjetas evaluadas

N° de Tarjeta	Utilidad Total	Utilidad Constante	Utilidad Tipo de Universidad	Utilidad Reputación Académica	Utilidad Calidad de la Enseñanza	Utilidad Infraestructura Académica	Utilidad Infraestructura Deportiva	Utilidad Perspectiva de Empleo
1	5.02143	4.55714	0.07143	0.42143	-1.27857	0.5	0.23571	0.51429
2	5.59285	4.55714	-0.07143	-0.42143	1.27857	-0.5	0.23571	0.51429
3	6.26429	4.55714	0.07143	-0.42143	1.27857	0.5	-0.23571	0.51429
4	3.40715	4.55714	-0.07143	0.42143	-1.27857	-0.5	-0.23571	0.51429
5	3.00713	4.55714	-0.07143	-0.42143	-1.27857	0.5	0.23571	-0.51429
6	5.54999	4.55714	0.07143	0.42143	1.27857	-0.5	0.23571	-0.51429
7	5.93571	4.55714	-0.07143	0.42143	1.27857	0.5	-0.23571	-0.51429
8	1.67857	4.55714	0.07143	-0.42143	-1.27857	-0.5	-0.23571	-0.51429

FUENTE: Elaboración propia

Las estimaciones obtenidas permiten realizar las siguientes aplicaciones más comunes del Análisis Conjunto: la segmentación de mercados (usando el comando **caSegmentation**) y la simulación de productos (usando el comando **ShowAllSimulations**). Las descripciones podemos observarla en la tabla (III.4).

### 3.11 APLICACIONES MÁS COMUNES DE LA REPRESENTACIÓN DE LA ESTRUCTURA DE PREFERENCIAS

#### 3.11.1 SEGMENTACIÓN

La segmentación permite identificar grupos de encuestados con preferencias similares. En el paquete Conjoint se cuenta con la función **caSegmentation** el cual permite realizar análisis de conglomerados dividiendo a los encuestados en n clusters usando el método k-means (función kmeans del paquete stats).

Por ejemplo, se buscó identificar dos grupos de estudiantes que se diferencien en las puntuaciones según sus preferencias universitarias, considerando que se hizo la aplicación a dos niveles distintos (4to. y 5to. sec.). Se utilizó la siguiente función:

Funciones R v. 3.1.1 – Segmentación

```
caSegmentation(y=prefe, x=code, c=2)
```

Figura III.31-Segmentación usando el método K-Means

```
K-means clustering with 2 clusters of sizes 18, 17

Cluster means:
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]      [,8]
1 6.173611 6.618056 7.048611 4.437500 4.048611 6.493056 7.118056 2.118056
2 3.801471 4.507353 5.433824 2.316176 1.904412 4.551471 4.683824 1.213235

Clustering vector:
[1] 1 1 1 1 1 1 1 2 2 2 1 2 2 1 2 1 1 2 1 1 2 2 2 1 2 2 1 2 2 2 1 1 2 1

Within cluster sum of squares by cluster:
[1] 333.5486 353.1324
(between_SS / total_SS = 29.1 %)

Available components:
[1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
[6] "betweenss"    "size"         "iter"         "ifault"
```

FUENTE: Resultados R v. 3.1.1

Una vez identificados los conglomerados se hizo un análisis agregado por cada grupo de modo que se comparen sus resultados.

El primer conglomerado contaba con 18 observaciones; mientras el segundo conglomerado con 17 observaciones. Luego, como una interpretación inicial, observando los resultados de los valores centrales por cada conglomerado (figura III.31) el primer grupo estaba constituido por aquellos que colocaron un mayor puntaje a todos los perfiles.

El análisis realizado al primer conglomerado (ver anexo 5) nos dio proporcionó los siguientes resultados:

- La estimación de los parámetros del modelo señalaban que el factor Tipo de Universidad no es significativa al 5%.
- Las estimaciones de las utilidades parciales indicaban que el atributo Tipo de universidad con el nivel Universidad nacional presentaba una mayor utilidad en comparación al nivel Universidad particular.
- El atributo más importante, según la figura (III.33), resultó ser la Calidad de la Enseñanza (32,56%), seguido de la Perspectiva de Empleo (15,85%), Infraestructura Académica (15,53%), Tipo de Universidad (14,07%), Reputación Académica (13,43%), y la Infraestructura Deportiva (8,57%).
- El coeficiente de determinación obtenido resultó con un  $R^2=0,53$ .

Figura III.32-Análisis del primer conglomerado

```

Call:
lm(formula = frml)

Residuals:
    Min       1Q   Median       3Q      Max
-4,4931 -1,1181 -0,1181  0,9514  4,8819

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   5,50694    0,13531   40,700 < 2e-16 ***
factor(x$Tipo)1  0,04861    0,13531    0,359  0,7199
factor(x$Repu)1  0,54861    0,13531   4,055 8,38e-05 ***
factor(x$Cali)1  1,31250    0,13531   9,700 < 2e-16 ***
factor(x$Acad)1  0,59028    0,13531   4,363 2,51e-05 ***
factor(x$Depo)1  0,32639    0,13531   2,412  0,0172 *
factor(x$Empl)1  0,56250    0,13531   4,157 5,65e-05 ***
---
Signif. codes:  0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

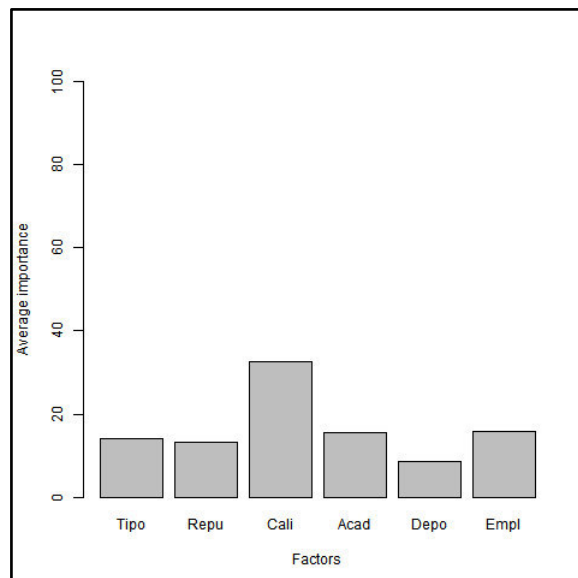
Residual standard error: 1,624 on 137 degrees of freedom
Multiple R-squared:  0.5273,    Adjusted R-squared:  0.5066
F-statistic: 25.47 on 6 and 137 DF,  p-value: < 2,2e-16

[1] "Part worths (utilities) of levels (model parameters for whole sample):"
      levnms      utls
1  intercept  5,5069
2  Nacional   0,0486
3  Particular -0,0486
4      Si     0,5486
5      No    -0,5486
6      Si     1,3125
7      No    -1,3125
8      Si     0,5903
9      No    -0,5903
10     Si     0,3264
11     No    -0,3264
12     Si     0,5625
13     No    -0,5625
[1] "Average importance of factors (attributes):"
[1] 14,07 13,43 32,56 15,53  8,57 15,85
[1] Sum of average importance: 100,01
[1] "Chart of average factors importance"

```

FUENTE: Resultados R v. 3.1.1

Figura III.33-Importancia relativa de los atributos en el primer conglomerado



FUENTE: Resultados R v. 3.1.1

Luego, el análisis realizado al segundo conglomerado (ver Anexo 6) nos proporcionaron los siguientes resultados:

- La estimación de los parámetros del modelo señalaban, que los factores Tipo de Universidad, Reputación Académica e Infraestructura Deportiva no son significativas al 5%.
- Las estimaciones de las utilidades parciales indicaban que en el atributo Tipo de universidad, el nivel Universidad Particular presentó una mayor utilidad en comparación al nivel Universidad Nacional. Diferente al primer conglomerado donde es el nivel Nacional el de mayor índice.
- El atributo más importante, según la figura (III.35), resultó ser la Calidad de la Enseñanza (31,54%), seguido de la Perspectiva de Empleo (17,23%), Reputación Académica (16,48%), Infraestructura Académica (14,72%), Tipo de Universidad (10,25%), y la Infraestructura Deportiva (9,79%).
- El coeficiente de determinación disminuyó en comparación al primer conglomerado ( $R^2=0,41$ ).



Figura III.34-Análisis del segundo conglomerado

```

Call:
lm(formula = frm1)

Residuals:
    Min       1Q   Median       3Q      Max
-3,5515 -0,9044 -0,2132  1,0956  4,4926

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    3,5515     0,1504   23,609 < 2e-16 ***
factor(x$Tipo)1 -0,1985     0,1504   -1,320  0,18926
factor(x$Repu)1  0,2868     0,1504    1,906  0,05884 .
factor(x$Cali)1  1,2426     0,1504    8,261  1,5e-13 ***
factor(x$Acad)1  0,4044     0,1504    2,688  0,00813 **
factor(x$Depo)1  0,1397     0,1504    0,929  0,35477
factor(x$Empl)1  0,4632     0,1504    3,079  0,00253 **
---
Signif. codes:  0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

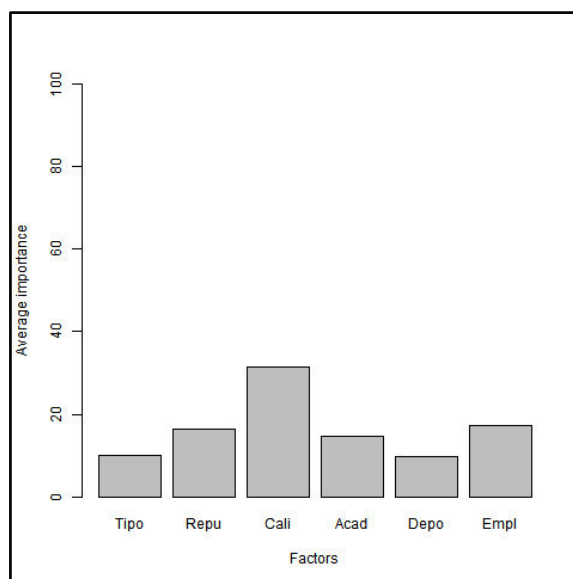
Residual standard error: 1,754 on 129 degrees of freedom
Multiple R-squared:  0.4141,    Adjusted R-squared:  0.3869
F-statistic: 15.2 on 6 and 129 DF,  p-value: 4,113e-13

[1] "Part worths (utilities) of levels (model parameters for whole sample):"
      levnms      utls
1  intercept  3,5515
2  Nacional -0,1985
3  Particular 0,1985
4      Si  0,2868
5      No -0,2868
6      Si  1,2426
7      No -1,2426
8      Si  0,4044
9      No -0,4044
10     Si  0,1397
11     No -0,1397
12     Si  0,4632
13     No -0,4632
[1] "Average importance of factors (attributes):"
[1] 10,25 16,48 31,54 14,72  9,79 17,23
[1] Sum of average importance: 100,01
[1] "Chart of average factors importance"

```

FUENTE: Resultados R v. 3.1.1

Figura III.35-Importancia relativa de los atributos en el segundo conglomerado



FUENTE: Resultados R v. 3.1.1

Por último, en la tabla (III.8) se presentan las combinaciones ideales obtenidos de las estimaciones en la evaluación general y en cada uno de los conglomerados.

Tabla III.8-Comparativo de perfiles ideales

ATRIBUTOS	GRUPO GENERAL	PRIMER CONGLOMERADO	SEGUNDO CONGLOMERADO
<b>TIPO DE UNIVERSIDAD</b>	Particular	Nacional	Particular
<b>REPUTACIÓN ACADÉMICA</b>	Si	Si	Si
<b>CALIDAD DE LA ENSEÑANZA</b>	Si	Si	Si
<b>INFRAESTRUCTURA ACADÉMICA</b>	Si	Si	Si
<b>INFRAESTRUCTURA DEPORTIVA</b>	Si	Si	Si
<b>PERSPECTIVA DE EMPLEO</b>	Si	Si	Si

FUENTE: Elaboración propia

### 3.11.2 SIMULACIÓN

Esta aplicación de la representación de la estructura de preferencias nos permite simular nuevas combinaciones, es decir nuevos productos. Con el objetivo de determinar las preferencias de estos. Valiéndose de los resultados obtenidos en el análisis principal.

Por ejemplo, en la figura (III.14) se presentó la matriz de perfiles codificados (“code”). Donde los valores considerados eran 1 ó 2. Para el atributo Tipo de universidad, el 1 representaba a la Universidad Nacional y el 2 a la Universidad Particular. En los demás atributos (Reputación académica, Calidad de la enseñanza, Infraestructura académica, Infraestructura deportiva y Perspectivas de empleo) el 1 representaba Si y el 2 No. Además, para la estimación general se consideraron sólo 8 perfiles de las 64 combinaciones posibles de los niveles de cada atributo.

Por lo tanto, la idea fue simular las puntuaciones de preferencias que presentarían los perfiles no considerados en el análisis principal.

En primer lugar se planteó las combinaciones, observado en la figura (III.36), del cual se deseaban obtener la simulación de sus preferencias. Estas combinaciones se las registro en una matriz denominada “sim”.

Figura III.36-Matriz de perfiles a simular

	Tipo	Repu	Cali	Acad	Depo	Empl
1	1	1	1	1	1	2
2	2	1	1	1	1	1
3	1	2	1	1	2	2
4	2	2	1	1	2	2
5	1	1	2	1	1	1
6	2	1	2	1	1	1
7	1	2	2	2	2	2

FUENTE: Resultados R v. 3.1.1

Si tenemos algunos perfiles de simulación de preferencias universitarias, se puede estimar la participación de los perfiles de simulación utilizando el modelo probabilístico BTL (Bradley-Terry-Luce) con la función `caBTL()`; el modelo logit con la función `caLogit ()` o el modelo de utilidad máxima (“first position”) con la función `caMaxUtility()`. Pero en lugar de estas tres funciones, es posible utilizar sólo una función de paquete Conjoint: la función `ShowAllSimulations()`.

Luego, a través del comando `ShowAllSimulations` del paquete Conjoint cuyos argumentos incluían a la matriz de los perfiles a simular (“sim”) observado en la figura (III.36), la matriz de preferencias universitarias (“prefem”) observado en el Anexo 2 y la matriz de perfiles codificados (“code”) presentado en la figura (III.14).

Funciones R v. 3.1.1 – Simulación de preferencias

```
ShowAllSimulations(sym=sim, y=prefem, x=code)
```

Figura III.37- Simulación de preferencias-porcentajes de participación

	TotalUtility	MaxUtility	BTLmodel	LogitModel
1	6,41	20,00	18,18	23,03
2	7,58	71,43	20,85	46,80
3	5,09	0,00	14,28	7,46
4	5,24	2,86	14,01	6,59
5	4,88	2,86	14,02	7,57
6	5,02	2,86	13,74	7,36
7	1,54	0,00	4,92	1,20

FUENTE: Resultados R v. 3.1.1

Se observa de los resultados obtenidos en la figura (III.37) que la mayor cuota de preferencia lo obtuvo el “producto 2”. Y la menor, como era de esperarse por los niveles de los atributos considerados, el “producto 7”.

## CONCLUSIONES

En el siguiente apartado se van a presentar las conclusiones del estudio:

- Los resultados, obtenidos en la aplicación del Análisis Conjunto para la “identificación de preferencias universitarias”, fueron los siguientes:
  - A nivel de la importancia relativa de cada atributo en la evaluación principal, los resultados indicaron que el atributo más valorado fue la Calidad de la Enseñanza (32,06%); por el contrario, el menos valorado fue la Infraestructura Deportiva (9,16%).
  - Todos los parámetros estimados del modelo son significativos al 5% (excepto el factor Tipo de Universidad), a nivel de un análisis agregado.
  - A nivel de las utilidades parciales de los niveles estimados indicaron que los niveles con mayor preferencia se presentaban en el Tipo de Universidad (Particular), Reputación Académica (Si), Calidad de la Enseñanza (Si), Infraestructura Académica (Si), Infraestructura Deportiva (Si) y Perspectiva de Empleo (Si).
  - La medición de la fiabilidad temporal se dio en base a la correlación de Pearson ( $r=0,97$ ). La evaluación de validación interna se obtuvo en función al coeficiente de determinación  $R^2$ . Haciendo un análisis agregado se obtuvo un  $R^2=0,39$ . De forma individual, considerando las preferencias de un estudiante, nos daba cuenta de un  $R^2=0,98$ . Así mismo, la evaluación de la capacidad predictiva del modelo señaló un coeficiente  $r=0,97$ .
- El análisis estadístico se realizó con el programa R en su versión 3.1.1. El paquete utilizado fue Conjoint (Andrzej Bak y Tomasz Bartlomowicz). El paquete cuenta con algoritmos específicamente para la evaluación del Análisis Conjunto tradicional,

basado en la metodología de recolección de datos perfiles completos. El uso de las funciones del package Conjoint no demandaron mayor complicación y permitieron, de manera detallada y completa, llevar a cabo el análisis de la técnica Análisis Conjunto tradicional de perfiles completos.

- Las aplicaciones de la representación de la estructura de preferencias se dieron a nivel de la segmentación y la simulación:
  - A nivel de la segmentación, se dividieron a los estudiantes en dos conglomerados, usando el método k-means. La diferencia más importante encontrada entre los dos grupos es que el primer grupo (conformado por 18 estudiantes) otorgó una mayor utilidad al nivel Universidad Nacional en comparación del nivel Universidad Particular en el atributo Tipo de Universidad. Mientras el grupo 2 (conformado por 17 estudiantes) fue lo contrario. Ambos grupos coincidieron en otorgar una mayor importancia relativa al atributo Calidad de la Enseñanza en comparación a los demás atributos. Por último, los coeficientes de determinación  $R^2$  mejoraron cuando se hizo el análisis por cada conglomerado en comparación a la evaluación de todos los estudiantes ( $R^2=0,53$  para el grupo 1 y  $R^2=0,41$  para el grupo 2).
  - A nivel de la simulación conjunta, se simularon preferencias de combinaciones no consideradas inicialmente. Obteniéndose porcentajes de participación valiéndose de los modelos probabilísticos BTL (Bradley-Terry-Luce), el modelo logit y el modelo de utilidad máxima.

## LISTA DE REFERENCIAS

- Bak, A., & Bartłomowicz, T. (2009). Conjoint analysis method and its implementation in conjoint R package. *Wroclaw: University of Economics*.
- Bak, A., & Bartłomowicz, T. (2012). Package conjoint. Conjoint analysis package.
- Bird, M., Kogan, L., Sánchez, L., & Villa, J. (2014). Hacia un nuevo modelo de las decisiones de compra: lógicas de consumo en jóvenes universitarios limeños.
- Cárdenas, E. (2006). *En busca del perfil del candidato político ganador mediante la aplicación del análisis conjunto Lima-Perú*. Universidad Nacional de Ingeniería. Programa Cybertesis PERÚ.
- Casielles, R. V. (1990). Investigación de las preferencias del consumidor mediante "análisis conjunto": importancia para el diseño de nuevos productos. *Documentos de trabajo (Universidad de Oviedo. Facultad de Ciencias Económicas)*(10), 1-30.
- Cortés, J. P. (2006). Análisis Conjunto. Ingeniería de Marketing IN58B. Departamento de Ingeniería Industrial. Universidad de Chile.
- Green, P., & Rao, V. (1971). Conjoint measurement for quantifying judgmental data. *Journal of Marketing research*, 355-363.
- Green, P., & Srinivasan, V. (1978). Conjoint analysis in consumer research: issues and outlook. *Journal of consumer research*, 103-123.
- Hair, J., Anderson, R., Tatham, R., & Black, W. (1999). *Análisis multivariante. 5ª edición*. editorial Prentice Hall: Madrid.
- Juez, P., & Díez, F. (1996). Capítulo 10 Análisis Conjunto. *Probabilidad y Estadística en Medicina: Aplicaciones en la Práctica Clínica y en la Gestión Sanitaria* (pp. 239-276).
- Lopes, S. D. F. (2011). Análisis conjunto: Teoría, campos de aplicación y conceptos inherentes. *Estudios y perspectivas en turismo*, 20(2), 341-366.
- Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of mathematical psychology*, 1(1), 1-27.
- Marquina, P., & Vásquez, A. (2013). Consumer social responses to CSR initiatives versus corporate abilities. *Journal of Consumer Marketing*, 100, 111.
- Moorhouse, J. (2012). What price university education?: Evidence from a conjoint analysis.
- Orme, B. (2002). Formulating attributes and levels in conjoint analysis. *Sawtooth Software Research Paper*, 1-4.
- Orme, B. (2003). Which conjoint method should I use? *Sawtooth Software Research Paper Series*.
- Orme, B. (2006). A short history of conjoint analysis. *Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research*.

- Papuico, K. (2015). Características estereotípicas que intervienen en el proceso de evaluación de candidatos electorales.
- Pérez, C. (2011). Capítulo 9 Análisis Conjunto. *Técnicas de segmentación: conceptos, herramientas y aplicaciones* (pp. 445-489): Garceta.
- Picón, E., Varela, J., & Braña, T. (2006). *Análisis conjunto*: Editorial La Muralla, S.A.
- Rao, V. R. (2014). *Applied conjoint analysis*: Springer.
- Salinas, J. (2012). *Estudio de preferencias usando el Análisis Conjunto*. Paper presented at the XIV Congreso Nacional de Estudiantes de Estadística, Universidad Nacional Mayor de San Marcos.
- Schnettler, B., Shene, C., Rubilar, M., Miranda, H., Sepúlveda, J., Denegri, M., & Lobos, G. (2010). Aceptación hacia yogurt con diferentes ingredientes funcionales en consumidores de supermercados del sur de Chile. *Arch Latinoam Nutr*, 60(4), 380-390.
- Soutar, G. N., & Turner, J. P. (2002). Students' preferences for university: a conjoint analysis. *International Journal of Educational Management*, 16(1), 40-45.
- Varela, J., Rial, A., & García, A. (2003). Capítulo 12 Análisis Conjunto. In J. Varela & J. Varela (Eds.), *Análisis multivariable para las Ciencias Sociales* (pp. 507-566). España.
- Varela, J., Rial, A., & Picón, E. (2014). *Tema 6 - Análisis Conjunto*. Universidade de Santiago de Compostela.
- Walsh, C., Barry, C., Dunnett, A., & Moorhouse, J. (2013). Choosing a university: the results of a longitudinal study using conjoint analysis.



# ANEXOS

## ANEXO 1 – INSTRUMENTO DE RECOLECCIÓN DE PREFERENCIAS

## IDENTIFICACIÓN DE PREFERENCIAS UNIVERSITARIAS

EDAD:  AÑOSMARQUE CON UNA X  
LO SIGUIENTENIVEL: 4TO. SEC. ☐  
5TO. SEC. ☐SEXO: FEMENINO ☐  
MASCULINO ☐

LEA DETENIDAMENTE LA DESCRIPCIÓN VERBAL DE CADA TARJETA

## TARJETA N° 1

## Estudio sobre Preferencias Universitarias

Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.

TIPO DE UNIVERSIDAD:	Particular
REPUTACIÓN ACADÉMICA:	Si
CALIDAD DE LA ENSEÑANZA:	No
INFRAESTRUCTURA ACADÉMICA:	Si
INFRAESTRUCTURA DEPORTIVA:	Si
PERSPECTIVAS DE EMPLEO:	Si

1	2	3	4	5	6	7	8	9	10
Muy poco preferida									Extremadamente preferida

## TARJETA N° 2

## Estudio sobre Preferencias Universitarias

Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.

TIPO DE UNIVERSIDAD:	Nacional
REPUTACIÓN ACADÉMICA:	No
CALIDAD DE LA ENSEÑANZA:	Si
INFRAESTRUCTURA ACADÉMICA:	No
INFRAESTRUCTURA DEPORTIVA:	Si
PERSPECTIVAS DE EMPLEO:	Si

1	2	3	4	5	6	7	8	9	10
Muy poco preferida									Extremadamente preferida

## TARJETA N° 3

## Estudio sobre Preferencias Universitarias

Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.

TIPO DE UNIVERSIDAD:	Particular
REPUTACIÓN ACADÉMICA:	No
CALIDAD DE LA ENSEÑANZA:	Si
INFRAESTRUCTURA ACADÉMICA:	Si
INFRAESTRUCTURA DEPORTIVA:	No
PERSPECTIVAS DE EMPLEO:	Si

1	2	3	4	5	6	7	8	9	10
Muy poco preferida									Extremadamente preferida

## TARJETA N° 4

## Estudio sobre Preferencias Universitarias

Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.

TIPO DE UNIVERSIDAD:	Nacional
REPUTACIÓN ACADÉMICA:	Si
CALIDAD DE LA ENSEÑANZA:	No
INFRAESTRUCTURA ACADÉMICA:	No
INFRAESTRUCTURA DEPORTIVA:	No
PERSPECTIVAS DE EMPLEO:	Si

1	2	3	4	5	6	7	8	9	10
Muy poco preferida									Extremadamente preferida

## TARJETA N° 5

## Estudio sobre Preferencias Universitarias

Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.

TIPO DE UNIVERSIDAD:	Nacional
REPUTACIÓN ACADÉMICA:	No
CALIDAD DE LA ENSEÑANZA:	No
INFRAESTRUCTURA ACADÉMICA:	Si
INFRAESTRUCTURA DEPORTIVA:	Si
PERSPECTIVAS DE EMPLEO:	No

1	2	3	4	5	6	7	8	9	10
Muy poco preferida									Extremadamente preferida

## TARJETA N° 6

## Estudio sobre Preferencias Universitarias

Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.

TIPO DE UNIVERSIDAD:	Particular
REPUTACIÓN ACADÉMICA:	Si
CALIDAD DE LA ENSEÑANZA:	Si
INFRAESTRUCTURA ACADÉMICA:	No
INFRAESTRUCTURA DEPORTIVA:	Si
PERSPECTIVAS DE EMPLEO:	No

1	2	3	4	5	6	7	8	9	10
Muy poco preferida									Extremadamente preferida

## TARJETA N° 7

## Estudio sobre Preferencias Universitarias

Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.

TIPO DE UNIVERSIDAD:	Nacional
REPUTACIÓN ACADÉMICA:	Si
CALIDAD DE LA ENSEÑANZA:	Si
INFRAESTRUCTURA ACADÉMICA:	Si
INFRAESTRUCTURA DEPORTIVA:	No
PERSPECTIVAS DE EMPLEO:	No

1	2	3	4	5	6	7	8	9	10
Muy poco preferida									Extremadamente preferida

## TARJETA N° 8

## Estudio sobre Preferencias Universitarias

Por favor, rodee con un círculo uno de los números que aparecen a continuación según sean sus preferencias por la siguiente alternativa.

TIPO DE UNIVERSIDAD:	Particular
REPUTACIÓN ACADÉMICA:	No
CALIDAD DE LA ENSEÑANZA:	No
INFRAESTRUCTURA ACADÉMICA:	No
INFRAESTRUCTURA DEPORTIVA:	No
PERSPECTIVAS DE EMPLEO:	No

1	2	3	4	5	6	7	8	9	10
Muy poco preferida									Extremadamente preferida

### ANEXO 2 – MATRIZ DE PREFERENCIAS UNIVERSITARIAS

	perfil_1	perfil_2	perfil_3	perfil_4	perfil_5	perfil_6	perfil_7	perfil_8
1	7	6	8	5	4	6	8	2
2	4	6	9	2	2	8	7	1
3	5	8	6	4	3	2	7	1
4	7	5	9	5	5	6	6	1
5	7	4	9	2	4	7	6	3
6	7	8	9	8	6	9	7	5
7	6	8	9	4	2	7	7	1
8	3	2	8	1	1	5	4	1
9	1	1	9	1	1	7	2	1
10	3	4	2	2	2	7	8	1
11	5	8	5	7	6	9	10	5
12	5	3	6	2	2	5	6	1
13	8	5	4	5	2	5	3	1
14	5	6	5	7	4	7	9	1
15	7	5	6	3	3	4	6	1
16	9	8	9	5	4	7	8	1
17	4	6	8	4	4	5	6	1
18	5	6	3	1	3	4	4	1
19	6	7	4	7	5	3	8	1
20	6	6	5	4	4	6	6	1
21	2	5	5	2	1	6	8	1
22	2	2	3	2	1	5	6	1
23	6	4	7	2	2	4	5	1
24	8	6	6	3	6	10	5	1
25	3	8	9	1	2	1	4	1
26	4	7	8	6	2	4	3	1
27	4	5	8	5	4	8	8	2
28	4	3	6	4	3	4	5	1
29	4	2	2	1	1	2	2	2
30	1	9	2	1	3	4	2	3
31	4	3	5	4	3	5	5	1
32	5	8	7	3	2	6	7	1
33	6	7	6	2	4	5	6	1
34	2	7	8	2	1	6	6	1
35	8	5	7	5	6	8	5	7

### ANEXO 3 – MATRIZ DE PERFILES DE RESERVA (HOLDOUT)

	perfil_1	perfil_2	perfil_3	perfil_4	perfil_5	perfil_6	perfil_7	perfil_8
1	4	8	8	2	1	9	10	1
2	5	7	6	2	1	9	7	1
3	4	6	7	2	3	3	6	1
4	5	5	7	3	2	2	4	1
5	2	3	1	1	4	6	2	1

### ANEXO 4 – MATRIZ DE PERFILES PARA EVALUAR LA FIABILIDAD TEMPORAL

	perfil_1	perfil_2	perfil_3	perfil_4	perfil_5	perfil_6	perfil_7	perfil_8
1	8	7	7	5	4	8	6	1
2	3	5	7	2	1	4	6	1
3	7	5	5	2	2	6	7	1
4	7	3	6	4	8	9	5	4
5	3	7	8	2	3	4	8	1
6	7	7	7	5	5	7	8	1
7	9	5	7	4	4	5	5	1
8	7	3	9	3	2	5	6	1
9	9	8	9	6	6	7	8	2
10	7	4	10	6	3	2	9	1

## ANEXO 5 – MATRIZ DE PREFERENCIAS. PRIMER CONGLOMERADO

	perfil_1	perfil_2	perfil_3	perfil_4	perfil_5	perfil_6	perfil_7	perfil_8
1	7	6	8	5	4	6	8	2
2	4	6	9	2	2	8	7	1
3	5	8	6	4	3	2	7	1
4	7	5	9	5	5	6	6	1
5	7	4	9	2	4	7	6	3
6	7	8	9	8	6	9	7	5
7	6	8	9	4	2	7	7	1
8	5	8	5	7	6	9	10	5
9	5	6	5	7	4	7	9	1
10	9	8	9	5	4	7	8	1
11	4	6	8	4	4	5	6	1
12	6	7	4	7	5	3	8	1
13	6	6	5	4	4	6	6	1
14	8	6	6	3	6	10	5	1
15	4	5	8	5	4	8	8	2
16	5	8	7	3	2	6	7	1
17	6	7	6	2	4	5	6	1
18	8	5	7	5	6	8	5	7

## ANEXO 6 – MATRIZ DE PREFERENCIAS. SEGUNDO CONGLOMERADO

	perfil_1	perfil_2	perfil_3	perfil_4	perfil_5	perfil_6	perfil_7	perfil_8
1	3	2	8	1	1	5	4	1
2	1	1	9	1	1	7	2	1
3	3	4	2	2	2	7	8	1
4	5	3	6	2	2	5	6	1
5	8	5	4	5	2	5	3	1
6	7	5	6	3	3	4	6	1
7	5	6	3	1	3	4	4	1
8	2	5	5	2	1	6	8	1
9	2	2	3	2	1	5	6	1
10	6	4	7	2	2	4	5	1
11	3	8	9	1	2	1	4	1
12	4	7	8	6	2	4	3	1
13	4	3	6	4	3	4	5	1
14	4	2	2	1	1	2	2	2
15	1	9	2	1	3	4	2	3
16	4	3	5	4	3	5	5	1
17	2	7	8	2	1	6	6	1

## ANEXO 7 – GLOSARIO DE TÉRMINOS

**ATRIBUTO.-** También denominado factor, es cada una de las características que definen a un producto o servicio (por ejemplo en un celular el color, el tamaño, la marca, el modelo, etc.).

**NIVEL.-** Es cada una de las alternativas u opciones que presenta un atributo (tomando el ejemplo anterior, el atributo color presenta los niveles blanco, negro, rojo, etc.)

**PATH-WORTH.-** Se denominan así a las utilidades parciales, los cuales son valores numéricos que expresan el grado de preferencia hacia un determinado nivel de un atributo. Estos se calculan como resultado de aplicar el Análisis Conjunto.

**IMPORTANCIA RELATIVA-** Son valores numéricos representados en porcentajes que nos indican las preferencias por cada atributo.

**PERFIL-** También se le denomina alternativa, estímulo o tratamiento. Es una combinación entre los niveles de los atributos en estudio. El cual nos da una idea del producto o servicio a elegir.

**ESCENARIO-** También se le denomina tarea. Es cada uno de los perfiles en el Análisis Conjunto Basado en la Elección donde un encuestado tiene que escoger el de mayor preferencia.

**DISEÑOS FACTORIALES FRACCIONADOS ORTOGONALES.-** Son diseños fraccionados que permiten seleccionar combinaciones de perfiles para estimar las utilidades con un número menor de perfiles del total de combinaciones posibles.

**ESTÍMULO HOLDOUT.-** Son perfiles de reserva. Estos perfiles son evaluados por los encuestados pero no participan en la estimación de utilidades. Se utiliza con fines de conocer la capacidad predictiva del modelo (validación cruzada).

## ANEXO N° 8

### DESARROLLO MATRICIAL DE LA ESTIMACIÓN OLS CON VARIABLES DUMMY

#### INTRODUCCIÓN

La estimación de utilidades parciales en el Análisis Conjunto se da de manera individual. Y los valores de las utilidades para toda la muestra se obtienen como resultado de promediar las utilidades parciales de cada individuo. Así también, los valores obtenidos de la importancia relativa de los atributos, utilidad total, etc.

En este anexo vamos a realizar el proceso de estimación de manera matricial. Recordemos que en el capítulo III hemos obtenido las estimaciones pero como resultado de la aplicación de las funciones del paquete **Conjoint** a través del programa estadístico R.

El desarrollo paso a paso de las estimaciones para el modelo conjunto permitirá un mejor entendimiento de los resultados obtenidos.

#### DESARROLLO

Para verificar como se desarrolló las estimaciones globales, como ya se ha señalado, el Análisis Conjunto realiza el proceso por cada individuo. Por lo tanto, los parámetros a estimar para el individuo  $i$  seguirán el criterio general de estimación:  $\hat{\beta} = (X'X)^{-1}X'Y$

En el desarrollo aplicativo se consideró al estudiante 2. Por lo tanto, lo primero que obtendremos es el vector de preferencias (Y) para este individuo. (Ver anexo 2, matriz de preferencias universitarias)

```
> Y
      V1
[1,]  4
[2,]  6
[3,]  9
[4,]  2
[5,]  2
[6,]  8
[7,]  7
[8,]  1
```

Luego, se consideró la matriz  $X$ . Al respecto hay que recordar que la estimación se realizó con el método de mínimos cuadrados para variables ficticias (dummy). Ya que el modelo de preferencia considera a los distintos niveles de los atributos como discretas. Por lo tanto, si el estudio consideró 6 atributos y cada atributo contaba con sólo 2 niveles al momento de diseñar esta matriz  $X$ , cada variable estuvo representada sólo por una columna. Considerando inicialmente una columna para la estimación del intercepto.

La codificación respondió en primer lugar a las combinaciones obtenidas del diseño factorial fraccionado ortogonal. Recordar que inicialmente se obtuvo 64 combinaciones ( $2^6=64$  seis atributos con dos niveles cada uno). Y luego se redujo a sólo 8 tratamientos, para un adecuado recojo de los datos a nivel de los encuestados.

Figura A.1-Perfiles obtenidos del Diseño Factorial Fraccionado Ortogonal

	Tipo	Repu	Cali	Acad	Depo	Empl
6	Particular	Si	No	Si	Si	Si
11	Nacional	No	Si	No	Si	Si
20	Particular	No	Si	Si	No	Si
29	Nacional	Si	No	No	No	Si
39	Nacional	No	No	Si	Si	No
42	Particular	Si	Si	No	Si	No
49	Nacional	Si	Si	Si	No	No
64	Particular	No	No	No	No	No

La forma de reparametrización de los distintos niveles de los atributos antes de proceder a la estimación de los parámetros se dio del siguiente modo:

- Tipo de universidad (Tipo) : Nacional (**1**) Particular(**-1**)
- Reputación académica (Repu) : Si (**1**) No(**-1**)
- Calidad de la enseñanza (Cali) : Si (**1**) No(**-1**)
- Infraestructura académica (Acad) : Si (**1**) No(**-1**)
- Infraestructura deportiva (Depo) : Si (**1**) No(**-1**)
- Perspectiva de empleo (Empl) : Si (**1**) No(**-1**)

Las matrices X y su transpuesta X' quedan determinadas del modo siguiente:

```
> X
      V1 V2 V3 V4 V5 V6 V7
[1,]  1 -1  1 -1  1  1  1
[2,]  1  1 -1  1 -1  1  1
[3,]  1 -1 -1  1  1 -1  1
[4,]  1  1  1 -1 -1 -1  1
[5,]  1  1 -1 -1  1  1 -1
[6,]  1 -1  1  1 -1  1 -1
[7,]  1  1  1  1  1 -1 -1
[8,]  1 -1 -1 -1 -1 -1 -1
```

```
> t(X)
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
V1      1      1      1      1      1      1      1      1
V2     -1      1     -1      1      1     -1      1     -1
V3      1     -1     -1      1     -1      1      1     -1
V4     -1      1      1     -1     -1      1      1     -1
V5      1     -1      1     -1      1     -1      1     -1
V6      1      1     -1     -1      1      1     -1     -1
V7      1      1      1      1     -1     -1     -1     -1
```

Luego, el producto  $X'X$



```
> t(X) %*% X
      V1 V2 V3 V4 V5 V6 V7
V1  8  0  0  0  0  0  0
V2  0  8  0  0  0  0  0
V3  0  0  8  0  0  0  0
V4  0  0  0  8  0  0  0
V5  0  0  0  0  8  0  0
V6  0  0  0  0  0  8  0
V7  0  0  0  0  0  0  8
```

Y la matriz inversa  $(X'X)^{-1}$


```
> solve(t(X) %*% X)
      V1      V2      V3      V4      V5      V6      V7
V1 0.125 0.000 0.000 0.000 0.000 0.000 0.000
V2 0.000 0.125 0.000 0.000 0.000 0.000 0.000
V3 0.000 0.000 0.125 0.000 0.000 0.000 0.000
V4 0.000 0.000 0.000 0.125 0.000 0.000 0.000
V5 0.000 0.000 0.000 0.000 0.125 0.000 0.000
V6 0.000 0.000 0.000 0.000 0.000 0.125 0.000
V7 0.000 0.000 0.000 0.000 0.000 0.000 0.125
```

Así mismo,  $X'Y$

```
> t(X) %*% Y
      V1
V1 39
V2 -5
V3 3
V4 21
V5 5
V6 1
V7 3
```

Por último, se obtuvo los valores estimados de las utilidades parciales  $\hat{\beta} = (X'X)^{-1}X'Y$

```
> solve(t(X) %*% X) %*% t(X) %*% Y
      V1
V1 4.875
V2 -0.625
V3 0.375
V4 2.625
V5 0.625
V6 0.125
V7 0.375
```



Estos valores obtenidos corresponden a las utilidades parciales del modelo conjunto para el segundo encuestado. El primer valor 4,875 es el parámetro del intercepto. De forma seguida se presentan las utilidades parciales de los atributos.

Los mismos resultados se obtienen utilizando la función **caModel** (package **Conjoint**) para el segundo encuestado. Por esta razón se considera el argumento **y=prefem[2,]** el cual me indica que toma sólo los valores de la fila dos de la matriz de preferencias que corresponden al segundo estudiante encuestado.



Para el análisis se utilizó la matriz de preferencias (**prefem**) (Anexo 2) y no el vector (**prefe**) (Figura III.13). Algo importante a señalar es que todas las funciones del paquete Conjoint trabajan sin problema con la matriz de preferencias (**prefem**); mientras, con el vector de preferencias (**prefe**) sólo puede ser utilizado en algunas de las funciones.

```
> caModel(y=prefem[2,], x=code)
```

Call:  
lm(formula = frml)

Residuals:

1	2	3	4	5	6	7	8
-0.375	-0.375	0.375	0.375	0.375	0.375	-0.375	-0.375

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	4.875	0.375	13.000	0.0489 *
factor(x\$Tipo)1	-0.625	0.375	-1.667	0.3440
factor(x\$Repu)1	0.375	0.375	1.000	0.5000
factor(x\$Cali)1	2.625	0.375	7.000	0.0903 .
factor(x\$Acad)1	0.625	0.375	1.667	0.3440
factor(x\$Depo)1	0.125	0.375	0.333	0.7952
factor(x\$Empl)1	0.375	0.375	1.000	0.5000

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.061 on 1 degrees of freedom  
Multiple R-squared: 0.9827, Adjusted R-squared: 0.8786  
F-statistic: 9.444 on 6 and 1 DF, p-value: 0.2441

Luego, para la obtención de las utilidades parciales de toda la muestra (esto es, considerando a los 35 encuestados) se toma el promedio de las utilidades parciales por cada individuo (utilidades medias).

Los valores de las utilidades parciales para cada uno de los encuestados se obtienen a través de la función **caPartUtilities**. Para referencia podemos apreciar en la siguiente tabla las estimaciones que se obtuvieron anteriormente para el segundo encuestado.

```
> caPartUtilities(y=prefe, x=code, z=level)
```

	intercept	Nacional	Particular	Si	No	Si	No	Si	No	Si	No	Si	No
[1,]	5.750	0.000	0.000	0.750	-0.750	1.250	-1.250	1.000	-1.000	0.000	0.000	0.750	-0.750
[2,]	4.875	-0.625	0.625	0.375	-0.375	2.625	-2.625	0.625	-0.625	0.125	-0.125	0.375	-0.375
[3,]	4.500	1.000	-1.000	0.000	0.000	1.250	-1.250	0.750	-0.750	0.000	0.000	1.250	-1.250
[4,]	5.500	-0.250	0.250	0.500	-0.500	1.000	-1.000	1.250	-1.250	0.250	-0.250	1.000	-1.000
[5,]	5.250	-1.250	1.250	0.250	-0.250	1.250	-1.250	1.250	-1.250	0.250	-0.250	0.250	-0.250
[6,]	7.375	-0.125	0.125	0.375	-0.375	0.875	-0.875	-0.125	0.125	0.125	-0.125	0.625	-0.625
[7,]	5.500	-0.250	0.250	0.500	-0.500	2.250	-2.250	0.500	-0.500	0.250	-0.250	1.250	-1.250
[8,]	3.125	-1.125	1.125	0.125	-0.125	1.625	-1.625	0.875	-0.875	-0.375	0.375	0.375	-0.375
[9,]	2.875	-1.625	1.625	-0.125	0.125	1.875	-1.875	0.375	-0.375	-0.375	0.375	0.125	-0.125
[10,]	3.625	0.375	-0.375	1.375	-1.375	1.625	-1.625	0.125	-0.125	0.375	-0.375	-0.875	0.875
[11,]	6.875	0.875	-0.875	0.875	-0.875	1.125	-1.125	-0.375	0.375	0.125	-0.125	-0.625	0.625
[12,]	3.750	-0.500	0.500	0.750	-0.750	1.250	-1.250	1.000	-1.000	0.000	0.000	0.250	-0.250
[13,]	4.125	-0.375	0.375	1.125	-1.125	0.125	-0.125	0.125	-0.125	0.875	-0.875	1.375	-1.375
[14,]	5.500	1.000	-1.000	1.500	-1.500	1.250	-1.250	0.250	-0.250	0.000	0.000	0.250	-0.250
[15,]	4.375	-0.125	0.125	0.625	-0.625	0.875	-0.875	1.125	-1.125	0.375	-0.375	0.875	-0.875
[16,]	6.375	-0.125	0.125	0.875	-0.875	1.625	-1.625	1.125	-1.125	0.625	-0.625	1.375	-1.375
[17,]	4.750	0.250	-0.250	0.000	0.000	1.500	-1.500	0.750	-0.750	0.000	0.000	0.750	-0.750
[18,]	3.375	0.125	-0.125	0.125	-0.125	0.875	-0.875	0.375	-0.375	1.125	-1.125	0.375	-0.375
[19,]	5.125	1.625	-1.625	0.875	-0.875	0.375	-0.375	0.625	-0.625	0.125	-0.125	0.875	-0.875
[20,]	4.750	0.250	-0.250	0.750	-0.750	1.000	-1.000	0.500	-0.500	0.750	-0.750	0.500	-0.500
[21,]	3.750	0.250	-0.250	0.750	-0.750	2.250	-2.250	0.250	-0.250	-0.250	0.250	-0.250	0.250
[22,]	2.750	0.000	0.000	1.000	-1.000	1.250	-1.250	0.250	-0.250	-0.250	0.250	-0.500	0.500
[23,]	3.875	-0.625	0.625	0.375	-0.375	1.125	-1.125	1.125	-1.125	0.125	-0.125	0.875	-0.875
[24,]	5.625	-0.625	0.625	0.875	-0.875	1.125	-1.125	0.625	-0.625	1.875	-1.875	0.125	-0.125
[25,]	3.625	0.125	-0.125	-1.375	1.375	1.875	-1.875	0.875	-0.875	-0.125	0.125	1.625	-1.625
[26,]	4.375	0.125	-0.125	-0.125	0.125	1.125	-1.125	-0.125	0.125	-0.125	0.125	1.875	-1.875
[27,]	5.500	0.000	0.000	0.750	-0.750	1.750	-1.750	0.500	-0.500	-0.250	0.250	0.000	0.000
[28,]	3.750	0.000	0.000	0.500	-0.500	0.750	-0.750	0.750	-0.750	-0.250	0.250	0.500	-0.500
[29,]	2.000	-0.500	0.500	0.250	-0.250	0.000	0.000	0.250	-0.250	0.250	-0.250	0.250	-0.250
[30,]	3.125	0.625	-0.625	-1.125	1.125	1.125	-1.125	-1.125	1.125	1.125	-1.125	0.125	-0.125
[31,]	3.750	0.000	0.000	0.750	-0.750	0.750	-0.750	0.500	-0.500	0.000	0.000	0.250	-0.250
[32,]	4.875	0.125	-0.125	0.375	-0.375	2.125	-2.125	0.375	-0.375	0.375	-0.375	0.875	-0.875
[33,]	4.625	0.125	-0.125	0.125	-0.125	1.375	-1.375	0.875	-0.875	0.875	-0.875	0.625	-0.625
[34,]	4.125	-0.125	0.125	-0.125	0.125	2.625	-2.625	0.125	-0.125	-0.125	0.125	0.625	-0.625
[35,]	6.375	-1.125	1.125	0.125	-0.125	-0.125	0.125	0.125	-0.125	0.375	-0.375	-0.125	0.125

Luego, si tomamos los promedios por cada columna obtendremos las utilidades medias. Que vendrían a ser las **estimaciones finales de los parámetros para el modelo conjunto**.

```
> colMeans(as.matrix(caPartUtilities(y=prefem, x=code, z=level)))
```

intercept	Nacional	Particular	Si	No	Si
4.55714286	-0.07142857	0.07142857	0.42142857	-0.42142857	1.27857143
No	Si	No	Si	No	Si
-1.27857143	0.50000000	-0.50000000	0.23571429	-0.23571429	0.51428571
No	-0.51428571				

Del mismo modo, para la obtención de las importancias relativas de los atributos, considerando a todos los encuestados, se obtienen en primer lugar las importancias relativas para cada uno de los individuos. Al respecto no hay una forma directa de obtener las importancias relativas por cada uno de los individuos en forma conjunta. Por lo tanto, se desarrolló un bucle **For** haciendo uso de la función **caImportance**.


```
vector=matrix(data=0,nr=35,nc=6)
for (i in 1:35)
{
  for (j in 1:6)
  {
    imp<-as.matrix(caImportance(y=prefem[i,], x=code))
    vector[i,j]=imp[j,1]
  }
}
vector
colMeans(vector)
```

Obteniéndose una matriz (vector) donde visualizamos las importancias relativas de los 6 atributos por cada uno de los encuestados. Y al final, se obtienen los promedios por cada columna los cuales me representan las importancias medias.

```

> vector
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]  0.00 20.00 33.33 26.67  0.00 20.00
[2,] 13.16  7.89 55.26 13.16  2.63  7.89
[3,] 23.53  0.00 29.41 17.65  0.00 29.41
[4,]  5.88 11.76 23.53 29.41  5.88 23.53
[5,] 27.78  5.56 27.78 27.78  5.56  5.56
[6,]  5.56 16.67 38.89  5.56  5.56 27.78
[7,]  5.00 10.00 45.00 10.00  5.00 25.00
[8,] 25.00  2.78 36.11 19.44  8.33  8.33
[9,] 36.11  2.78 41.67  8.33  8.33  2.78
[10,]  7.89 28.95 34.21  2.63  7.89 18.42
[11,] 21.87 21.88 28.12  9.37  3.13 15.62
[12,] 13.33 20.00 33.33 26.67  0.00  6.67
[13,]  9.38 28.12  3.12  3.13 21.88 34.37
[14,] 23.53 35.29 29.41  5.88  0.00  5.88
[15,]  3.13 15.63 21.87 28.12  9.38 21.87
[16,]  2.17 15.22 28.26 19.57 10.87 23.91
[17,]  7.69  0.00 46.15 23.08  0.00 23.08
[18,]  4.17  4.17 29.17 12.50 37.50 12.50
[19,] 36.11 19.44  8.33 13.89  2.78 19.44
[20,]  6.67 20.00 26.67 13.33 20.00 13.33
[21,]  6.25 18.75 56.25  6.25  6.25  6.25
[22,]  0.00 30.77 38.46  7.69  7.69 15.38
[23,] 14.71  8.82 26.47 26.47  2.94 20.59
[24,] 11.90 16.67 21.43 11.90 35.71  2.38
[25,]  2.08 22.92 31.25 14.58  2.08 27.08
[26,]  3.57  3.57 32.14  3.57  3.57 53.57
[27,]  0.00 23.08 53.85 15.38  7.69  0.00
[28,]  0.00 18.18 27.27 27.27  9.09 18.18
[29,] 33.33 16.67  0.00 16.67 16.67 16.67
[30,] 11.90 21.43 21.43 21.43 21.43  2.38
[31,]  0.00 33.33 33.33 22.22  0.00 11.11
[32,]  2.94  8.82 50.00  8.82  8.82 20.59
[33,]  3.12  3.13 34.38 21.87 21.88 15.63
[34,]  3.33  3.33 70.00  3.33  3.33 16.67
[35,] 56.25  6.25  6.25  6.25 18.75  6.25
> colMeans(vector)
[1] 12.209714 14.910286 32.060857 15.139143  9.160571 16.517143

```




Las importancias medias de las importancias obtenidas se verifican con los valores de las importancias encontrados a través de la función **caImportance** para toda la muestra.

```

> imp<-caImportance(y=prefe, x=code)
> print("Importance summary: ", quote=FALSE)
[1] Importance summary:
> print(imp)
[1] 12.21 14.91 32.06 15.14  9.16 16.52
> print(paste("Sum: ", sum(imp)), quote=FALSE)
[1] Sum: 100

```



Así mismo, las estimaciones de las utilidades totales para cada encuestado se obtienen con la función **caTotalUtilities**



```
> caTotalUtilities(y=prefe, x=code)
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
[1,] 7.000 6.000 8.000 5.000 4.000 6.000 8.000 2.000
[2,] 4.375 6.375 8.625 1.625 1.625 7.625 7.375 1.375
[3,] 4.250 7.250 6.750 4.750 3.750 2.750 6.250 0.250
[4,] 7.750 5.750 8.250 4.250 4.250 5.250 6.750 1.750
[5,] 7.250 4.250 8.750 1.750 3.750 6.750 6.250 3.250
[6,] 7.625 8.625 8.375 7.375 5.375 8.375 7.625 5.625
[7,] 6.000 8.000 9.000 4.000 2.000 7.000 7.000 1.000
[8,] 3.625 2.625 7.375 0.375 0.375 4.375 4.625 1.625
[9,] 2.625 2.625 7.375 -0.625 -0.625 5.375 3.625 2.625
[10,] 2.625 3.625 2.375 2.375 2.375 7.375 7.625 0.625
[11,] 4.875 7.875 5.125 7.125 6.125 9.125 9.875 4.875
[12,] 5.000 3.000 6.000 2.000 2.000 5.000 6.000 1.000
[13,] 7.875 4.875 4.125 5.125 2.125 5.125 2.875 0.875
[14,] 5.250 6.250 4.750 6.750 3.750 6.750 9.250 1.250
[15,] 6.625 4.625 6.375 3.375 3.375 4.375 5.625 0.625
[16,] 8.875 7.875 9.125 5.125 4.125 7.125 7.875 0.875
[17,] 4.500 6.500 7.500 3.500 3.500 4.500 6.500 1.500
[18,] 4.375 5.375 3.625 1.625 3.625 4.625 3.375 0.375
[19,] 5.625 6.625 4.375 7.375 5.375 3.375 7.625 0.625
[20,] 6.000 6.000 5.000 4.000 4.000 6.000 6.000 1.000
[21,] 1.750 4.750 5.250 2.250 1.250 6.250 7.750 0.750
[22,] 2.000 2.000 3.000 2.000 1.000 5.000 6.000 1.000
[23,] 5.875 3.875 7.125 2.125 2.125 4.125 4.875 0.875
[24,] 8.625 6.625 5.375 2.375 5.375 9.375 5.625 1.625
[25,] 2.625 7.625 9.375 1.375 2.375 1.375 3.625 0.625
[26,] 4.625 7.625 7.375 5.375 1.375 3.375 3.625 1.625
[27,] 4.750 5.750 7.250 4.250 3.250 7.250 8.750 2.750
[28,] 4.500 3.500 5.500 3.500 2.500 3.500 5.500 1.500
[29,] 3.500 1.500 2.500 1.500 1.500 2.500 1.500 1.500
[30,] 0.375 8.375 2.625 1.625 3.625 4.625 1.375 2.375
[31,] 4.500 3.500 4.500 3.500 2.500 4.500 5.500 1.500
[32,] 4.625 7.625 7.375 3.375 2.375 6.375 6.625 0.625
[33,] 5.625 6.625 6.375 2.375 4.375 5.375 5.625 0.625
[34,] 2.125 7.125 7.875 1.875 0.875 5.875 6.125 1.125
[35,] 8.125 5.125 6.875 4.875 5.875 7.875 5.125 7.125
```

Luego, si consideramos las medias de cada columna (ocho perfiles), se obtendrán las utilidades totales medias. Estos valores nos indican cuál ha sido el perfil (tarjeta) de mayor preferencia.

```
> colMeans(caTotalUtilities(y=prefe, x=code))
[1] 5.021429 5.592857 6.264286 3.407143 3.007143 5.550000 5.935714 1.678571
```

Los resultados nos indican que la combinación en la tarjeta n° 3 presentó la mayor preferencia; de forma contraria, la tarjeta n° 8 fue la de menos preferencia.

Todos los resultados obtenidos anteriormente lo obtenemos a través del uso de la función: **Conjoint**

```
> Conjoint(y=prefe, x=code, z=level)

Call:
lm(formula = frml)

Residuals:
    Min       1Q   Median       3Q      Max
-4,5929 -1,0214 -0,0214  1,4500  5,3214

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    4,55714    0,11668   39,058 < 2e-16 ***
factor(x$Tipo)1 -0,07143    0,11668   -0,612  0,540915
factor(x$Repu)1  0,42143    0,11668    3,612  0,000362 ***
factor(x$Cali)1  1,27857    0,11668   10,958 < 2e-16 ***
factor(x$Acad)1  0,50000    0,11668    4,285  2,53e-05 ***
factor(x$Depo)1  0,23571    0,11668    2,020  0,044333 *
factor(x$Empl)1  0,51429    0,11668    4,408  1,50e-05 ***
---
Signif. codes:  0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Residual standard error: 1,952 on 273 degrees of freedom
Multiple R-squared:  0.3911,    Adjusted R-squared:  0.3778
F-statistic: 29.23 on 6 and 273 DF,  p-value: < 2,2e-16

[1] "Part worths (utilities) of levels (model parameters for whole sample):"
  levnm      utls
1 intercept  4,5571
2 Nacional -0,0714
3 Particular 0,0714
4 Si 0,4214
5 No -0,4214
6 Si 1,2786
7 No -1,2786
8 Si 0,5
9 No -0,5
10 Si 0,2357
11 No -0,2357
12 Si 0,5143
13 No -0,5143

[1] "Average importance of factors (attributes):"
[1] 12,21 14,91 32,06 15,14 9,16 16,52
[1] Sum of average importance: 100
```